Misallocation Under Trade Liberalization

Yan Bai — University of Rochester, NBER

Keyu Jin — LSE, CEPR

Dan Lu * — The Chinese University of Hong Kong

April 19, 2020

Abstract

This paper formalizes a classic idea that in second-best environments trade can induce welfare losses. In a framework that incorporates distortion wedges into a Melitz model, we analyze a channel in which trade can reduce allocative efficiency arising from the reallocation of resources. A key aggregate statistics that captures this negative selection is the gap between input and output shares. We derive sufficient conditions for welfare loss due to trade under important distributions. Using Chinese manufacturing data for the period 1998-2007, we show that welfare gains and productivity have qualitatively and quantitatively large departures from those predicted by standard models of trade.

Keywords: Capital and labor wedges, misallocation, trade liberalization, gains from trade, industrial policy

JEL classification: E23 F12 F14 F63 L25 O47

*Bai: Department of Economics, University of Rochester; email: yanbai06@gmail.com. Jin: Department of Economics, London School of Economics; email: k.jin@lse.ac.uk. Lu: Department of Economics, The Chinese University of Hong Kong; email: danlv27@gmail.com.

Acknowledgments: We thank George Alessandria, Mark Bils, Chang Tai Hsieh, Oleg Itskhoki, Narayana Kocherlakota, Ezra Oberfield, Michael Song, Kjetil Storesletten, Heiwai Tang, Fabrizio Zillibotti, Daniel Xu for their invaluable comments, as well as seminar participants in NBER Summer Institute, NBER China meeting, Tsinghua growth and institute, AWSOME conference of McMaster U., Philadelphia Fed Trade Workshop, St. Louis Fed Macro-Trade Conference, IMF, Fed Board, Chinese U. of Hong Kong, U. of Hong Kong, HongKong U. of Science and Technology, SHUFE, Fudan U. and NSE of Peking University. Keyu Jin thanks ESRC award P004253/1.

1 Introduction

The question of how much developing countries benefit from opening up to goods trade is a time-honoured subject, both of practical import and intellectual interest. Much has been understood about the nature and type of gains to trade, thanks to the remarkable progress made in the field of international trade in recent decades. Less clear, however, is why certain developing countries have benefited from trade more than others, and why certain countries have seemingly not benefitted much at all.[1] New trade theories suggest that developing countries have the most to gain from trade: if trade liberalization can induce reallocation of resources from less to more productive firms, aggregate productivity and welfare will rise in turn.

But developing countries are different in another respect: they are also subject to prevalent policy and institutional distortions. Examples include taxes and subsidies to certain firms, implicit guarantees and bailouts, preferential access to land and capital, and industrial policy and export promotion policies—common themes in developing countries. In the case of China, for instance, this explains why inefficient but politically-connected private firms and state-owned companies (SOE) have survived and even thrived. Implicit and explicit support for these firms combined with limited exit mechanisms for many SOEs have weakened firm selection effects, the upshot of which is a drag on aggregate productivity. Many believe that joining the WTO can potentially alleviate some of these problems by inviting direct competition from abroad.

How effective is this mechanism? Can trade necessarily improve allocations? Does trade necessarily lead to welfare gains for developing countries? These issues are far from obvious as alluded to by Rodríguez-Clare (2018), " [a] complication that may matter for the computation of the gains from trade is the presence of domestic distortions." This argument that trade may exert a different impact in a second-best environment has been an old age question posed by Bhagwati and Ramaswami (1963). Even in classic textbook analysis, there are discussions on the "domestic market failure argument against trade", that " [when] the theory of second best [is applied] to trade policy..., imperfections in the

[1] For example, Waugh (2010) shows, in large sample of countries, that poor countries do not systematically gain more from trade.

internal function of an economy may justify interfering in its external economic relations" (Krugman, Obstfeld, and Melitz (2015)). We formalize these ideas in the context of the new trade model variety. A main contribution in this paper is to derive a general theoretical welfare formula to analyze additional channels of trade in a second-best environment. A second contribution is to take advantage of firm-level data to gauge how much these effects matter.

Our modelling framework allows for firm-specific distortions into a two-country Melitz model. There are two dimensions of heterogeneity at the firm-level: productivity and distortions. These distortions are assumed to be exogenous output wedges or factor wedges. They drive differences in the marginal products across firms. We deliberately do not take a stand on where these distortions come from as various kinds of policy and institutional distortions are legion in developing countries. In this context, contrary to the mechanism underpinning the Melitz (2003) model and its extensions, i.e. that trade can induce a reallocation of resources from low productivity to high productivity firms, the presence of distortions can bring about the opposite and exacerbate misallocation. The reason is simple: distortions (for instance, tax and subsidies) act as a veil to a firm's true productivity. A firm may be producing in the market not because it is inherently productive, but because it is sufficiently subsidized. A mass of highly-subsidized but not adequately productive firms will export and expand at the cost of other more productive firms. The high productivity/ high tax firms which were marginally able to survive in the domestic market would be driven out as the other firms gain market share and drive up costs. In other words, the selection effect which brings about efficiency gains in the Melitz-type model is no longer based solely on productivity; it is determined jointly by firm productivity and distortions. Trade may thus *lower* the average productivity of firms.

To formalize this argument, we derive a general welfare formula that relates to canonical trade models, such as Arkolakis, Costinot, and Rodríguez-Clare (2012)) and Melitz and Redding (2015). We show that a key statistic in capturing this negative reallocation channel in the aggregate is the gap between aggregate input share in producing domestic goods and aggregate expenditure share on domestic goods. If the required inputs used for producing export is greater than the output share it yields, then the reduction in allocative efficiency arising from a reallocation of resources, occasioned by trade, can bring about a welfare loss.

The general formula for welfare nests several important cases. The special case in which there is only heterogeneity in productivity, which follows a Pareto distribution, yields the well-known result of ACR (after Arkolakis, Costinot, and Rodríguez-Clare (2012)). Under a more general distribution, the expression corresponds to that in Melitz and Redding (2015). In these cases, there is always a positive selection mechanism associated with trade. The second special case is one in which there is only heterogeneity in distortions, Pareto distributed, and this yields an analogue to ACR under misallocation. In this case, the selection is purely driven by distortion, and welfare is *always* lower in an open compared to a closed economy. Intuitively, under homogenous productivity, the efficient allocation should be firms have identical market shares. However, with distortions, the relatively subsidized firms produce more than in the efficient case, with the dispersion of sales (employment) reflecting the distortions. When opening up to trade, it is the relatively subsidized firms that export, in turn making the firm's distribution even more skewed—reflecting a worsening of allocation. Hence, there is always welfare loss. In this case, exporters are highly subsidized, and the required inputs used for producing exports is greater than the output share it yields. In fact, the gap between the two is a sufficient statistic for welfare loss. We also derive sufficient conditions for welfare losses under more general distribution functions.

In the general case where productivity and distortions coexist, they are competing in their impact on firm selection. The relative strength of the two depends on their joint distribution, and micro-level information still matters for welfare. The impact of allocative efficiency arising from a reallocation of resources is captured by the input-sale gap, along with additional structural microeconomic parameters. To our knowledge, the paper is the first to theoretically characterize welfare loss to trade, and derive sufficient conditions for welfare. The decomposition of welfare into a 'pure technology effect' and a 'resource reallocation effect' in this instance resonates with the decomposition in the recent work of Baqaee and Farhi (2020). Where as they focus on network effects, we focus on how distortions determine the welfare impact of trade.

The second contribution is to operationalize our results in the context of China. We choose China because it is an economy saddled with distortions, and one that recently experienced an important trade liberalization event. We use our structural model combined with micro data from Chinese manufacturing to conduct a quantitative analysis of the impact of trade on welfare and aggregate productivity. The main goal is to see how much departure there is from standard trade models that do not take into account pre-existing domestic distortions. We run counterfactual experiments for local changes in trade cost, as well as counterfactual experiments for domestic reforms. Our main conclusion is that welfare gains are much smaller when taking into account distortions; that there is a TFP loss of 3% as opposed to a TFP gain of 13.3% in the case without distortions, and that allocative inefficiency can induce a welfare loss of 18%. Welfare gains are half of what standard models yield.

To further investigate the key mechanisms implied by our model, we conduct out-of-sample tests by examining the differential patterns among exporters and non-exporters in both the cross-section and the time series dimension, in Section 3.4. These patterns fit broadly with predictions in the model.

It is important to point out that in the quantitative analyses we do not use directly empirically-measured wedges, observed correlations, or distributions in the data to assess the impact of trade on welfare. The reason is that the *observed* statistics are not the *underlying* ones: existing firms have been subject to selection and thus their observed distributions are not the true ones. The same reasoning goes for the observed correlation between productivity and wedges. As we show in Section 3.1, the presence of fixed costs and/or firm selection can drive a positive relationship between the two. For these reasons, the approach adopted in the quantitative exercises is to estimate the underlying joint distribution of wedges and productivity, costs of producing and exporting so as to match the observed patterns of firms' outputs, inputs, and exports. On this basis, we evaluate how the presence of distortions change the impact of trade on productivity and welfare, and how much trade has contributed to Chinese growth in a decomposition exercise. This contrasts with the reduced-form approach adopted in Berthou, Chung, Manova, and Bragard (2018), which uses empirically measured revenue productivity to assess the impact of trade reforms on aggregate productivity under misallocation.[2] Our works are broadly

[2] Berthou et al. (2018) theoretically and empirically assess the impact of trade reforms for 14 European countries and 20 industries over the period 1998-2011. They find that trade reforms have ambiguous effects

complementary, as we focus on theoretical analyses and a structural approach to inferring welfare gains for China, whilst they focus on an empirical assessment of trade on aggregate productivities for 14 European countries.

Costa-Scottini (2018) and Ho (2010) and Berthou et al. (2018) are other papers that introduce firm level distortions to trade models. Theoretically, Costa-Scottini (2018) and Ho (2010) assume a log linear relationship, and hence perfect correlation of (log) productivity and (log) wedges. This assumption of perfect correlation is both limiting in its scope of analysis, and also inconsonant with patterns in the data. Overall, the main departure from these papers is both theoretical and quantitative. Theoretically, we explicitly explore the selection and reallocation effects, while formalizing the efficiency loss amounting to a few statistic— in particular– the wedge between the input and output shares (and the related hazard function). We also provide sufficient conditions for welfare loss under specific distributions. From an empirical perspective, all three papers measure firms wedges or productivity directly from the data using the well-known 'TFPR', while we take a different approach. We avoid employing directly empirically-measured wedges, observed correlations, or distributions in the data to assess the impact of trade on welfare. Again, we demonstrate that 'TFPR' and 'TFPQ' in these models do not measure directly productivity and wedges, as they are affected by wedges, productivity *and* fixed costs; in addition, the observed joint distribution is the ex-post one after selection.

What makes our paper different from the important works of Hsieh and Klenow (2009), Baily, Hulten, and Campbell (1992), Restuccia and Rogerson (2008), Bartelsman, Haltiwanger, and Scarpetta (2009) is first of all, the open economy nature of our model, and secondly, the endogenous mechanism of entry/exit and the attendant firm selection effect. For our purposes, selection is vital. Trade affects resource allocation through an endogenous selection of firms. Furthermore, in our work, 'misallocation of resources' goes beyond the observed misallocation among a set of operating firms. Because policy distortions also act as a barrier to entry (and exit), there is also misallocation among potential entrants and incumbents– firms that should have entered the market in an efficient economy that couldn't, and firms that should have otherwise exited but have not. This reallocation along

on measured revenue productivity in the theory, while they are positive in the data.

the entry/exit margin can also be significant. Empirical works have also demonstrated the importance of entry and exit for China's growth.[3]

China is well suited for the study for multiple reasons: for its prevalent State interventions and policies; that a body of work has shown (see references below) that idiosyncratic distortions explain the majority of the dispersion in marginal products; and that trade liberalization has been an important recent phenomenon. Distortions in China manifest themselves in the form of substantial privileges of state owned enterprises over private firms, of connected private firms, or of firms belonging to particular locations. Specific policies that can drive these wedges include implicit subsidies such as soft budget constraints, favorable costs of capital, preferential tax treatments and implicit guarantees. Firms with political connections having access to special deals and receiving substantial benefits are also widely documented (see Guo, Jiang, Kim, and Xu (2013) and Bai, Hsieh, and Song (2019)). Wu (2018) conducts an empirical analysis and finds that policy distortions can be explained by investment promoting programs that favor such firms.

There is also substantial evidence coming from a number of papers that these idiosyncratic firm-distortions account for a large part of the observed dispersion in marginal products across firms in China. In principle, misallocation can arise from a variety of factors; but different approaches to disentangle them have come to similar conclusions that policy distortions are elemental.[4]

Many of these distortions are also presumably unrelated to trade. For instance, firms such as the car manufacturing company Chery have enjoyed easy access to land and capital from their local governments. Foxconn, the world's largest electronics contracter manufacturer, has enjoyed substantial tax breaks from many provinces including industrial land at significantly discounted prices. Tesla has recently received free land and subsidies from

[3] Brandt, Van Biesebroeck, and Zhang (2012) find that net entry accounts for roughly half of Chinese manufacturing productivity growth. The creation and selection of new firms in China's non-state sector has been particularly important.

[4] These factors could be technological frictions, such as adjustment costs, information frictions, financial frictions, or markups. Wu (2018) finds that policies account for the majority of the observed misallocation of capital, as opposed to financial frictions. Using a different approach and modeling framework, David and Venkateswaran (2017) find also that firm-specific distortions, rather than technological or information frictions, account for the majority of the observed dispersions in marginal products. Bai, Lu, and Tian (2018) disciplines financial frictions with firms' financing patterns, sales distribution and change of capital. They find that financial frictions cannot explain the observed relation between firms' measured distortions and size.

the local government of Shanghai. A recent study by Chen and Kung (2018) demonstrate the firms that are connected with political elites were able to obtain land at 80 to 90 percent discount over the period 2004-2016.

For the reasons above, the baseline model in the paper is to focus on domestic policy distortions. Still, one can ask whether some of the large dispersion of marginal products reflects endogenous distortions—those that can potentially change with trade liberalization. As a robustness check we examine a model of endogenous distortions with variable markup, and ask whether trade can mitigate these distortions and the misallocation of resources. Section 4.1 takes up a variable markup model. We show that these models 1) yield some obvious counterfactual predictions on the relationship between exporters and wedges; 2) that markup alone also explains little of the dispersion in wedges. To match the observed correlation and dispersion one would still need to include exogenous distortions. Moreover, the attendant pro-competitive effects in a model with endogenous markup may be 'elusive' as pointed out by Arkolakis, Costinot, Donaldson, and Rodríguez-Clare (2018).[5]

In this framework, positive firm selection is the central driving force for gains to trade. As such, it abstracts from other types of gains to trade, such as trade-induced technological diffusion (Alvarez, Buera, and Lucas Jr (2013) and Buera and Oberfield (2016)), adoption (Perla, Tonetti, and Waugh (2015) and Sampson (2015)) and innovation (Atkeson and Burstein (2010)). While these mechanisms in principle work to increase the gains to trade, with its quantitative significance a subject to debate,[6] it does not detract from the fact that the distortionary impact on allocation efficiency still induces large welfare losses, which is what we are interested in. Of course, distortions can also interact with some of these additional channels. For instance, in a model with firm innovation, one would need to consider the fact that distortions not only affect production decisions, but potentially also innovation decisions. These considerations go beyond the scope of this paper but deserve

[5]This paper makes the point that when more productive firms expand at the expense of less productive ones, thanks to trade, the aggregate markup tends to rise. Thus, overall, trade models with endogenous markups do not necessarily generate higher gains from trade.

[6]Perla, Tonetti, and Waugh (2015) and Atkeson and Burstein (2010), for instance, find that trade gains are not too different from ACR gains. In Perla, Tonetti, and Waugh (2015), there is trade-induced within-firm productivity improvements. However, their aggregate growth effects come with costs—losses in variety and reallocation of resources away from goods production. Thus, the aggregate effect on welfare is similar to ACR gains. Atkeson and Burstein (2010) show that general equilibrium effects limits the first-order effects on aggregate productivity even when there is firm-level innovation.

further consideration. We also do not consider how trade can reduce domestic distortions, for example if concurrent domestic reforms are requisite for joining the WTO or if quotas are removed (see Khandelwal, Schott, and Wei (2013)). As a robustness check, however, we allow for firms to face a different distribution of distortions when they start to export and examine welfare and efficiency gains therein.

Taken together, our quantitative analysis is meant to highlight the first-order effects of a particular channel—allocative inefficiency, and also to compare it with benchmark results in the workhorse models of international trade. It is, however, not a comprehensive analysis of trade gains in the case of China. A key message of this paper is that in order for developing countries to reap the full gains of trade, simultaneous or antecedent domestic reforms aimed at reducing policy distortions may be crucial. The policy implication drawn from this framework is consistent with works indicating that policies aimed to neutralize domestic distortions may be complementary to trade liberalization (Chang, Kaltani, and Loayza (2009) and Harrison and Rodríguez-Clare (2010)). It counters other claims that trade liberalization should take precedence owing to positive firm selection (Asturias, Hur, Kehoe, and Ruhl (2016)).[7]

In sum, this paper shows that experiences of trade liberalization in developing countries should not be considered to be independent of micro-level distortions to which they are subject. Our paper demonstrates that the presence of policy distortions have a first-order quantitative effect on the gains to trade. The organization of the paper is as follows: Section 2 derives a theoretical framework of trade gains under misallocation. Section 3 provides a quantitative assessment on the impact of trade liberalization under misallocation, with various extensions of the benchmark framework. Section 4 discusses a model with endogenous distortions. Section 5 concludes.

[7]They show that the best sequence of reforms is to first decrease trade costs, then to improve contract enforcement, and, finally, to decrease the cost of firm creation.The reason is that an increase in competition leads to an expansion of productive firms and crowding out of less efficient ones. By liberalizing international trade first so as to impose firm selection early, inefficient firms are prevented from entering later when contract enforcement and firm entry costs are reformed. In contrast, we show that the selection mechanism is substantially weakened in the presence of distortions.

2 Theoretical Framework

2.1 Baseline Model

The world consists of two large open economies, Home and Foreign, with heterogenous firms. The two economies can differ in the size of labor and distribution of firms. Labor is immobile across countries and inelastic in supply.

Consumers. A representative consumer in the Home country chooses the amount of final goods C in order to maximize utility $u(C)$, subject to the budget constraint

$$PC = wL + \Pi + T,$$

where P is the price of final goods, L is labor, w is wage rate, Π is dividend income, and T is the amount of lump-sum transfers received from the government.

Final Goods Producers. Final goods producers are perfectly competitive, and combine intermediate goods using a CES production function

$$Q = \left[\int_{\omega \in \Omega} q(\omega)^{\frac{\sigma-1}{\sigma}} d\omega \right]^{\frac{\sigma}{\sigma-1}},$$

where σ is the elasticity of substitution across intermediate goods, and Ω is the endogenous set of goods. The corresponding final goods price index is thus

$$P = \left[\int_{\omega \in \Omega} p(\omega)^{1-\sigma} d\omega \right]^{\frac{1}{1-\sigma}},$$

where $p(\omega)$ is the price of good ω in the market. The individual demand for this good is thus given by

$$q(\omega) = \frac{p(\omega)^{-\sigma}}{P^{-\sigma}} Q.$$

Intermediate Goods Producers. There is a competitive fringe of potential entrants (in both countries) that can enter by paying a sunk entry cost of f_e units of labor. Potential entrants face uncertainty about their productivity in the industry. They also face a stochastic revenue wedge τ, which can be seen as a tax (>1) or subsidy (<1) on every revenue earned.[8] Once the sunk entry cost is paid, a firm draws its productivity φ and τ from a joint distribution, $g(\varphi, \tau)$ over $\varphi \in (0, \infty), \tau \in (0, \infty)$.[9]

Firms are monopolistically competitive. Production of each intermediate good entails fixed production cost of f units of labor and a constant variable cost that depends on firm productivity. The total labor required to produce $q(\varphi)$ units of a variety is therefore: [10]

$$\ell = f + \frac{q}{\varphi}.$$

(φ, τ) are idiosyncratic and independent across firms. The existence of a fixed production cost means that only a subset of firms produces—those that draw a sufficiently low productivity or high wedge cannot generate enough variable profits to cover the fixed production cost. If firms decide to export, they face a fixed exporting cost of f_x units of labor and iceberg variable costs of trade τ_x, which is greater than 1. Firms with the same productivity and distortion behave identically, and thus we can index firms by their (φ, τ) combination.

An intermediate goods firm thus solves the following problem

$$\max_{p,q} \ \frac{pq}{\tau} - \frac{w}{\varphi}q - wf \tag{1}$$

subject to the demand function $q = \frac{p^{-\sigma}}{P^{-\sigma}}Q$, henceforward suppressing ω for convenience. From here it is clear that a revenue tax is equivalent to a tax on all input costs incurred by the firm.

Firms are infinitely small, and thus take the aggregate price index as given. Equating the after-tax marginal revenue with marginal costs yields the standard result that equilibrium

[8]It is equivalent to an input wedge on all the input a firm uses.

[9]The model equilibrium is equivalent to a stationary equilibrium of a model allowing for constant exogenous probability of death δ and entry cost f_e/δ.

[10]We can easily extend the production to include capital, i.e. $k^{\alpha}\ell^{1-\alpha}$. The unit cost for producing q or fixed cost is $\alpha^{-\alpha}(1-\alpha)^{\alpha-1}w^{1-\alpha}r_k^{\alpha}$ where r_k is the rental cost of capital. In our model, we introduce one heterogeneous distortions at the firm level, and our τ is an output distortion, but it includes all input distortions that increase the marginal products of capital and labor by the same proportion as an output distortion. In the data, there are distortions that affect both capital and labor and distortions that change the marginal product of one of the factors relative to the other. In our quantitative exercises, we include both capital and labor, and the distortions on both factors.

prices are a mark-up over marginal costs:

$$p = \frac{\sigma}{\sigma - 1} \frac{w\tau}{\varphi}. \tag{2}$$

Optimal profits are then

$$\pi = \sigma^{-\sigma} (\sigma - 1)^{\sigma - 1} P^{\sigma} Q \tau^{-\sigma} w^{1-\sigma} \varphi^{\sigma - 1} - wf. \tag{3}$$

It immediately follows that given the fixed cost of production, there is a zero-profit cutoff productivity below which firms would choose not to produce, and exit the market. Thus, a firm would choose to produce only if $\varphi \geq \varphi^*(\tau)$. This cutoff productivity level satisfies

$$\varphi^*(\tau) = \frac{\sigma^{\frac{\sigma}{\sigma-1}}}{\sigma - 1} \left[\frac{wf}{P^{\sigma} Q} \right]^{\frac{1}{\sigma-1}} w \tau^{\frac{\sigma}{\sigma-1}}. \tag{4}$$

The cutoff productivity is now a function of the firm-specific distortion, and differs across firms facing different levels of distortions. Firms with a higher tax τ will have a higher cutoff for productivity. This means that low productivity firms that would have been otherwise excluded from the market can now enter the market and survive if sufficiently subsidized.

With trade, firms now have the option of exporting abroad. If a Home firm exports to the Foreign economy, it solves the following problem

$$\max \frac{p_x q_f}{\tau} - \frac{w}{\varphi} \tau_x q_f - w f_x$$

subject to the Foreign demand function $q = \frac{p_x^{-\sigma}}{P_f^{-\sigma}} Q_f$, where P_f and Q_f denote the aggregate price index and demand in Foreign. Given the same constant elasticity of demand in the domestic and export markets, equilibrium prices in the export market are a constant multiple of those in the domestic market:

$$p_x(\varphi, \tau) = \frac{\sigma}{\sigma - 1} \frac{w \tau_x \tau}{\varphi},$$

The optimal profit from servicing the Foreign market,

$$\pi_x = \sigma^{-\sigma}(\sigma-1)^{\sigma-1} P_f^{\sigma} Q_f \tau^{-\sigma} (w\tau_x)^{1-\sigma} \varphi^{\sigma-1} - w f_x, \tag{5}$$

yields an optimal cutoff for exporting:

$$\varphi_x^*(\tau) = \frac{\sigma^{\frac{\sigma}{\sigma-1}}}{\sigma-1} \left[\frac{w f_x \tau_x^{\sigma-1}}{P_f^{\sigma} Q_f} \right]^{\frac{1}{\sigma-1}} w \tau^{\frac{\sigma}{\sigma-1}}. \tag{6}$$

Consumer love of variety, a fixed production cost and additional fixed cost of exporting, mean that firms would never export without also selling in the domestic market. There are, hence, two cutoff productivities relevant for the domestic economy: one for entering the domestic market as given by (4) and one for entering the Foreign market, as given by (6). To the extent that taxes τ are constant across firms, the ratio $\varphi_x^*(\tau)/\varphi^*(\tau)$ is a constant and is greater than 1 so long as $\frac{\tau_x^{\sigma-1} f_x}{f} \frac{P^{\sigma} Q}{P_f^{\sigma} Q_f} > 1$. Analogously, firms in the Foreign country, which draw their productivity from a distribution $g_f(\varphi, \tau)$, are subject to two cutoff productivities, one for servicing their domestic market, and one for exporting to the Home economy

$$\varphi_f^*(\tau) = \frac{\sigma^{\frac{\sigma}{\sigma-1}}}{\sigma-1} \left[\frac{w_f f}{P_f^{\sigma} Q_f} \right]^{\frac{1}{\sigma-1}} w_f \tau^{\frac{\sigma}{\sigma-1}}, \tag{7}$$

$$\varphi_{xf}^*(\tau) = \frac{\sigma^{\frac{\sigma}{\sigma-1}}}{\sigma-1} \left[\frac{w_f f_x \tau_x^{\sigma-1}}{P^{\sigma} Q} \right]^{\frac{1}{\sigma-1}} w_f \tau^{\frac{\sigma}{\sigma-1}}. \tag{8}$$

where w_f denotes the Foreign wages, and the fixed cost of producing and exporting are assumed to be identical in the two economies. The government's budget is balanced so that

$$T = \int_{\omega \in \Sigma} \left(1 - \frac{1}{\tau}\right) p(\omega) q(\omega) d\omega,$$

where Σ is the endogenous set of home production.

The equilibrium features a constant mass of firms entering M_e and producing M, along with a ex-post distributions of productivity and distortion among operational firms $\mu(\varphi, \tau)$. The ex-post distribution $\mu(\varphi, \tau)$ is a truncation of the ex-ante productivity-distortion distribution, $g(\varphi, \tau)$, at the zero-profit cutoff productivity given by Eq.4:

$$\mu(\varphi, \tau) = \frac{g(\varphi, \tau)}{\int \int_{\varphi^*(\tau)}^{\infty} g(\varphi, \tau)\, d\varphi d\tau} \tag{9}$$

if $\varphi \geq \varphi^*(\tau)$; and $\mu(\varphi, \tau) = 0$ otherwise. The denominator is the probability of successful entry, denoted as

$$\omega_e = \int \int_{\varphi^*(\tau)}^{\infty} g(\varphi, \tau)\, d\varphi d\tau. \tag{10}$$

In equilibrium, the measure of producing firms equals the product of measure of entrants and the probability of entering, i.e.

$$\omega_e M_e = M.$$

We define the probability of exporting conditional on entry as

$$\omega_x = \int \int_{\varphi_x^*(\tau)}^{\infty} \mu(\varphi, \tau)\, d\varphi d\tau = \frac{\int \int_{\varphi_x^*(\tau)}^{\infty} g(\varphi, \tau)\, d\varphi d\tau}{\int \int_{\varphi^*(\tau)}^{\infty} g(\varphi, \tau)\, d\varphi d\tau}.$$

In an equilibrium with positive entry, the free entry condition requires that

$$\int \int_{\varphi^*(\tau)} \pi(\varphi, \tau) g(\varphi, \tau) d\varphi d\tau + \int \int_{\varphi_x^*(\tau)} \pi_x(\varphi, \tau) g(\varphi, \tau) d\varphi d\tau = w f_e. \tag{11}$$

The first term is the expected profits from domestic sales conditional on entry, multiplied by the probability of entry. The second term is the expect profits from export sales conditional on exporting, multiplied by the probability of exporting. The free entry condition requires that their sum be equal to the entry costs (in terms of labor).

The free entry condition (11), combined with optimal profit functions (3) and (5) gives an expression for the price index P:

$$P = \frac{\sigma}{\sigma - 1} \left[M \int \int_{\varphi^*(\tau)}^{\infty} \left(\frac{w\tau}{\varphi} \right)^{1-\sigma} \mu(\varphi, \tau)\, d\varphi d\tau + M_f \int \int_{\varphi_{xf}^*(\tau)}^{\infty} \left(\frac{w_f \tau_x \tau}{\varphi} \right)^{1-\sigma} \mu_f(\varphi, \tau)\, d\varphi d\tau \right]^{\frac{1}{1-\sigma}}, \tag{12}$$

where M and M_f denote the measure of operating firms in Home and Foreign. The Foreign price index P_f takes a similar form.

Goods market clearing. The assumption of a balanced trade results in

$$P_f^{\sigma}Q_f M \int\int_{\varphi_x^*(\tau)}^{\infty} \left(\frac{w\tau_x\tau}{\varphi}\right)^{1-\sigma} \mu\left(\varphi,\tau\right)d\varphi d\tau = P^{\sigma}QM_f \int\int_{\varphi_{xf}^*(\tau)}^{\infty} \left(\frac{w_f\tau_x\tau}{\varphi}\right)^{1-\sigma} \mu_f\left(\varphi,\tau\right)d\varphi d\tau. \tag{13}$$

Labor market clearing. In the equilibrium, the labor market condition yields

$$M = \frac{L}{\sigma\left(\frac{f_e}{\omega_e} + f + \omega_x f_x\right)}. \tag{14}$$

Normalizing the Home country wage rate to 1, there are eleven equations, the zero cutoff productivities for domestic production and exporting (4), (6), and its Foreign counterparts, the free entry conditions (11) along with its Foreign counterpart, the definition of the Home and Foreign price indices (12), and a goods market clearing/balanced trade equation (13), along with the measure of firms (14) and its Foreign counterpart. These equations yield the equilibrium consisting of eleven unknowns $\{\varphi^*(\tau),\ \varphi_x^*(\tau),\ \varphi_f^*(\tau),\ \varphi_{fx}^*(\tau),\ P,\ P_f,\ Q,\ Q_f,\ w_f,\ M,\ M_f\}$. A detailed derivation of the model is provided in Appendix A.

Proposition 1. *The allocations, entrants, and cutoff functions* $\{Q,\ Q_f,\ M,\ M_f,\ \varphi^*(\tau),\ \varphi_f^*(\tau),\ \varphi_x^*(\tau),\ \varphi_{xf}^*(\tau)\}$ *are independent of mean wedge* $\bar{\tau}$. *Prices* $\{P, P_f, w_f\}$ *change proportionally with* $\bar{\tau}, \bar{\tau}^f$, *i.e.* $P(\bar{\tau}_1)/P(\bar{\tau}_2) = \bar{\tau}_1/\bar{\tau}_2$, *and similarly for* P_f *and* w_f.

The proof of the proposition is straightforward. The proposition shows that increasing the mean of the wedges doesn't affect real variables. Hence misallocation of resources is not because of the average wedge across firms but heterogenous wedges.

2.2 Theoretical Comparative Static

We proceed to analyze welfare and efficiency with distortions. Welfare, or consumption, is given by:[11]

$$W = \frac{\sigma-1}{\sigma}\left[M_e \int\int_{\varphi^*(\tau)} \left(\varphi\frac{\overline{MRPL}}{MRPL_\tau}\right)^{\sigma-1} g(\varphi,\tau)d\varphi d\tau + M_e\frac{P_f^{\sigma}Q_f}{P^{\sigma}Q}\int\int_{\varphi_x^*(\tau)} \left(\frac{\varphi}{\tau_x}\frac{\overline{MRPL}}{MRPL_\tau}\right)^{\sigma-1} g(\varphi,\tau)d\varphi d\tau\right]^{\frac{1}{\sigma-1}}, \tag{15}$$

[11]Home's welfare depends on foreign exports, which are rewritten by using the balanced trade condition.

where $MRPL_\tau$ denotes the firm-specific marginal revenue product of labor, $MRPL_\tau = w\tau$, and $\overline{MRPL}$ denote the economy-wide marginal revenue product of labor, $\overline{MRPL} = PQ/L$. Hence, welfare is related to a weighted firm productivity, where the relative distortions are the weights. In an efficient case without distortions, all firms have the same marginal revenue product, $MRPL_\tau = \overline{MRPL}$ for any τ. Hence, Equation (15) shows that the source of welfare loss in the presence of firm-level distortions can arise from a misallocation of resources, captured by dispersions in $\overline{MRPL}/MRPL_\tau$, and a misallocation caused by selection and entry mechanisms captured by M_e, φ^*, φ_x^* being different from their respective efficient levels. Change in trade cost affects the economy through selection, entry and misallocation.

To understand how trade costs affect welfare through changes in misallocation (which includes selection and entry), we first present an expression for welfare in the presence of shocks to exogenous productivity—which can be compared with Baqaee and Farhi (2020) (henceforward BF) in the closed economy. The same reasoning can be applied to a trade liberalization episode. In subsequent analyses, this expression is further fleshed out to provide a formulation comparable to those in standard trade models.

Some definitions are in order: let $p_i q_i$ and ℓ_i be the total sales and variable labor of firm with productivity φ_i, λ the share of expenditure on domestic goods, and S the share of variable labor used in producing domestic goods.[12] Define $\hat{\varphi} = \varphi\tau^{\frac{\sigma}{1-\sigma}}$, so that the firm's log operating profit is proportional to $\log(\hat{\varphi})$. Recall that a firm produces if and only if its productivity is large enough or its wedge τ is small enough, i.e., $\varphi \geq \varphi^*(\tau) = \frac{\sigma^{\frac{\sigma}{\sigma-1}}}{\sigma-1}\left[\frac{wf}{P^\sigma Q}\right]^{\frac{1}{\sigma-1}} w\tau^{\frac{\sigma}{\sigma-1}}$. Thus, the production cutoff can be rewritten as $\hat{\varphi}^* = \frac{\sigma^{\frac{\sigma}{\sigma-1}}}{\sigma-1}\left[\frac{wf}{P^\sigma Q}\right]^{\frac{1}{\sigma-1}} w$, and a firm produces if and only if the combination of productivity and wedge satisfies

[12] Definition of the domestic expenditure share:

$$\lambda = \frac{\int\int_{\varphi^*(\tau)} \left(\frac{\varphi}{\tau}\right)^{\sigma-1} g(\varphi,\tau)d\varphi d\tau}{\int\int_{\varphi^*(\tau)} \left(\frac{\varphi}{\tau}\right)^{\sigma-1} g(\varphi,\tau)d\varphi d\tau + \frac{P_f^\sigma Q_f}{P^\sigma Q}\tau_x^{1-\sigma}\int\int_{\varphi_x^*(\tau)} \left(\frac{\varphi}{\tau}\right)^{\sigma-1} g(\varphi,\tau)d\varphi d\tau}.$$

Definition of the share of variable labor used in producing domestic goods:

$$S = \frac{\int\int_{\varphi^*(\tau)} \varphi^{\sigma-1}\tau^{-\sigma} g(\varphi,\tau)d\varphi d\tau}{\int\int_{\varphi^*(\tau)} \varphi^{\sigma-1}\tau^{-\sigma} g(\varphi,\tau)d\varphi d\tau + \frac{P_f^\sigma Q_f}{P^\sigma Q}\tau_x^{1-\sigma}\int\int_{\varphi_x^*(\tau)} \varphi^{\sigma-1}\tau^{-\sigma} g(\varphi,\tau)d\varphi d\tau}.$$

$\hat{\varphi} \geq \hat{\varphi}^*$.

There are two important elasticities, γ_λ—the elasticity of the cumulative sales share within the domestic market for firms above the domestic cutoff, and γ_s—the elasticity of the cumulative domestic (variable) labor share for firms above the domestic cutoff, both with respect to the cutoff $\hat{\varphi}$. These elasticities are also known as the hazard functions.[13]

Proposition 2. *In the presence of distortions,*

1. In a closed economy, the change in welfare associated with an exogenous productivity shock to firms with productivity φ_i is

$$dlnW = -d\ln P + (\sigma - 1)\left[\frac{p_i q_i}{PQ} - \frac{\ell_i}{L}\right] d\ln\varphi_i + \Big[\gamma_s(\hat{\varphi}^*) - \gamma_\lambda(\hat{\varphi}^*)\Big] d\ln\hat{\varphi}^*.$$

2. In an open economy, the change in welfare associated with an exogenous iceberg cost shock is

$$d\ln W = -d\ln P + \Big[-d\ln\lambda + d\ln S\Big] + \Big[\gamma_s(\hat{\varphi}^*) - \gamma_\lambda(\hat{\varphi}^*)\Big] d\ln\hat{\varphi}^*.$$

PROOF: Appendix B.1.

Technology shocks can have two effects on welfare, in a closed or an open economy case. The first is through a change in the aggregate price index P. A positive productivity shock or a negative trade shock lowers P and leads to a welfare gain. The second effect is coming from a change in the resources going into each firm. If a firm is relatively subsidized, its labor share is larger than its sales share. Thus, its expansion impinges negatively on welfare. In this closed economy, a firm's positive productivity shock may not raise aggregate productivity. This has a similar flavor to findings in Baqaee and Farhi

[13]Define $O(\hat{\varphi})$ as the cumulative sales share under any $\hat{\varphi}$ in the domestic market,

$$O(\hat{\varphi}) = \frac{\int\int_0^{\hat{\varphi}\tau^{\sigma/(\sigma-1)}} \varphi^{\sigma-1}\tau^{1-\sigma} g(\varphi,\tau) d\varphi d\tau}{\int\int_0^{\infty} \varphi^{\sigma-1}\tau^{1-\sigma} g(\varphi,\tau) d\varphi d\tau},$$

and $I(\hat{\varphi})$ the cumulative variable input share in the domestic market,

$$I(\hat{\varphi}) = \frac{\int\int_0^{\hat{\varphi}\tau^{\sigma/(\sigma-1)}} \varphi^{\sigma-1}\tau^{-\sigma} g(\varphi,\tau) d\varphi d\tau}{\int\int_0^{\infty} \varphi^{\sigma-1}\tau^{-\sigma} g(\varphi,\tau) d\varphi d\tau}.$$

The elasticity $\gamma_s(\hat{\varphi})$ is also proportional to the distribution of firm after-tax sales within the domestic market and is given by $\gamma_s(\hat{\varphi}) = -\frac{d\ln(1-I(\hat{\varphi}))}{d\ln\hat{\varphi}}$. The elasticity $\gamma_\lambda(\hat{\varphi})$ is given by $\gamma_\lambda(\hat{\varphi}) = -\frac{d\ln(1-O(\hat{\varphi}))}{d\hat{\varphi}}$.

(2020), which decomposes the macroeconomic impact of microeconomic shocks to productivity and wedges, and show that the effect on output can be decomposed into a "pure technology effect" and a "resource allocation effect". The former is the change in output holding fixed the share of resources going to each user; the latter is the change in output resulting from the reallocation of shares of resources across users. The difference between BF and our closed-economy model is the endogenous firm selection and entry. If the selection mechanism is shut off in our model, then the aggregate price index $-d\ln P$ becomes $p_i q_i/(PQ)$, which then translates into a formula for welfare change: $\frac{p_i q_i}{PQ} + (\sigma - 1)[\frac{p_i q_i}{PQ} - \frac{\ell_i}{L}]$. This expression says that the aggregate impact of firm i's productivity shock depends on its sales share and the gap between its sales share and labor share. Holding fixed the labor used by each firm, the productivity shock increases the producer's sales. But the shock also changes relative prices, and in turn demand, which causes a reallocation of labor among firms. If firm i is relatively subsidized, then $1/\tau_i$ is larger than the average level of distortion, and its labor share is larger than its output share. This producer is too large relative to the efficient allocation, and thus, reallocating labor towards this firm worsens allocative efficiency.

The same reasoning applies for trade cost shocks in an open economy. A shock to trade costs induces an endogenous selection of firms that induces a shuffling of resources. As relatively subsidized firms expand thanks to this 'technology' shock, allocative efficiency deteriorates. Thus, apart from a first-order welfare gain working through lowering the price index, a reduction in trade costs also incurs a first-order loss through the reallocation of resources.

This open-economy case is more complex compared to the closed-economy setting, for the reason that trade has a differential impact on firms. Some firms that remain domestic producers are not directly affected by the trade cost shock, while others may be selected into exporting. Or still yet, some firms may be ousted from producing altogether. Despite these heterogeneous effects, a neat result arises: the gap between aggregate input share and aggregate sales share is informative of the allocative efficiency. If the change in aggregate domestic labor share S is greater than the change in domestic expenditure share λ, then the trade cost shock could be welfare-reducing. More precisely, we can view the reallocation effect as arising from an intensive margin: holding selection fixed, i.e., $d\ln\hat{\varphi}^* = 0$, the shifting of resources among an existing set of firms is captured by $(dlnS - d\ln\lambda)$. From an extensive margin, where selection and cutoffs $\hat{\varphi}^*$ change, the gap of S and λ and γ_λ and γ_s summarize the reallocation of resources towards or away from more distorted firms (more on this below).

For our purposes, selection is vital. Trade affects resource allocation through an endogenous selection of firms. Even in a closed economy, selection and entry both affect the allocation of resources and welfare. This makes our exercise different from HK, apart from the open-economy nature of this model. The number of varieties is determined by entry and the probability of successful operating. Entry can be important because varieties affect welfare, and entry directly affects the number of varieties. Selection and entry make a difference also can be seen from that the special case of BF and HK, where productivities and wedges are jointly log-normal, the joint distribution is irrelevant—only the marginal distribution of the wedges matters. This is no longer the case when selection and entry are taken into account, as in our framework.

We now go one step further in understanding these reallocation effects, and how the welfare expression in Proposition 2 relates to trade gains arising from canonical trade models. We derive a general expression for changes in welfare associated with changes in trade costs. Arkolakis, Costinot, and Rodríguez-Clare (2012) (henceforward ACR) demonstrate that in the absence of distortions, welfare changes across a wide class of models can be inferred using two variables: (i) changes in the share of expenditure on domestic goods; and (ii) the elasticity of bilateral imports with respect to variable trade costs (the trade elasticity). Different trade models can have different micro-level predictions, sources of welfare gains, and different structural interpretations of the trade elasticity. But conditional on observed trade flows and an estimated trade elasticity, the welfare predictions are the same. The generality of this formulation, however, relies on a certain set of macro-level restrictions. Melitz and Redding (2015) (henceforth MR) show that under more general distribution functions for productivity, the trade elasticity is no longer invariant to trade costs and across markets, and therefore no longer a sufficient statistic for welfare. Micro-level information is still important for welfare.

In the analysis below, we first consider a fall in trade costs in an open economy equilibrium. The following proposition provides a general representation of welfare:

Proposition 3. *(**General Welfare Expression**) The change in welfare associated with an iceberg cost shock is*

$$
\begin{aligned}
d\ln W = \frac{1}{\gamma_s+\sigma-1}\Big[& -d\ln\lambda \qquad (ACR) \\
& + d\ln M_e \qquad (MR) \\
& + \frac{\sigma}{\sigma-1}(\gamma_s-\gamma_\lambda)d\ln M_e \\
& - \left(\sigma-1+\frac{\sigma\gamma_s}{\sigma-1}\right)d\ln\lambda + \left(\sigma-1+\frac{\sigma\gamma_\lambda}{\sigma-1}\right)d\ln S\Big] \qquad (Reallocation)
\end{aligned} \tag{16}
$$

PROOF: Appendix B.2.

The above proposition encapsulates welfare results for three different cases:

1. Without domestic distortions, $S=\lambda$ and $\gamma_s=\gamma_\lambda$. If productivity follows a Pareto distribution with parameter θ, $\gamma_\lambda=\theta-\sigma+1$ and $dlnM_e=0$. Hence,

$$dlnW=\frac{1}{\theta}\left[-dln\lambda\right]$$

as in ACR.

2. Under a general distribution function and without domestic distortions, $S=\lambda$, $\gamma_s=\gamma_\lambda$ (non constants), and $dlnM_e\neq 0$. Hence,

$$dlnW=\frac{1}{\gamma_\lambda+\sigma-1}\left[-dln\lambda+dlnM_e\right].$$

Here, the micro structure matters for γ_λ and hence welfare, as in MR.

3. With homogenous productivity and Pareto-distributed domestic distortion $1/\tau$ with parameter θ, $\gamma_\lambda=\frac{\sigma-1}{\sigma}(\theta-\sigma+1)$ and $\gamma_s=\frac{\sigma-1}{\sigma}(\theta-\sigma)$. Hence,

$$dlnW=\frac{\sigma}{\sigma-1}\left[dlnS-dln\lambda\right].$$

In this general welfare representation, the first term is referred to as *ACR*, the second and the third term relate to *entry* (which captures the changes to M_e), and the fourth term is brought about by distortions, and is referred to as a *reallocation* effect. Information on the change of domestic shares (sales and variable labor), the measure of entrants, the joint distribution of firms sales and variable inputs, and the cutoff firms (from which we know γ_λ and γ_s) are sufficient for computing the associated welfare change for a local change in trade cost.

It's useful to examine the special cases embedded therein. If there is only heterogeneity in productivity, Pareto-distributed (case 1), the ACR formula is recovered. The case without distortion and under a more general productivity distribution gives rise to MR. The second special case is that under misallocation, where there is only heterogeneity in distortions which is Pareto distributed, an analogue formula to ACR can be obtained: the difference in the change in the domestic labor and sales share provides a sufficient statistics for welfare.

In the two special cases (1) and (3), firm selection is either driven solely by productivity, or solely by distortions. The former implies that there is always a welfare improvement when the economy opens up to trade, whereas the latter implies that there is an unambiguous loss (see corollary below). In the more general case, productivity and distortions jointly determine firm selection. The resource reallocation is both one amongst existing firms (last term), and along the entry dimension (third term). Without distortions, a firm's share of input is equal to its share of sales, so that in aggregate, $S = \lambda$. In the presence of distortions, the two are no longer equal. The gap between input and sales shares is informative about changes in allocative efficiency. If the change in required inputs exceeds the change in revenue it produces, i.e. $dlnS < dln\lambda$, it means with further opening up, the input share used to produce for exports exceeds the export revenue share. Resources reallocation has induced an efficiency loss.[14]

Corollary 1. *(**Welfare Loss**) Under homogenous productivity and Pareto-distributed domestic*

[14] We also study the welfare expression in the Ricardian model and Armington model with distortions in Bai, Jin, and Lu (2020). The gap between domestic input share and sales share in total expenditure is still informative of resource reallocation. A proposition similar to Corollary 1 holds for these two models with distortions.

distortion $1/\tau$ with parameter θ,

$$dlnW = \frac{\sigma}{\sigma - 1}\left[dlnS - dln\lambda\right],$$

and

1. *Moving from a closed economy to an open economy always entails a welfare loss, as $\lambda > S$.*
2. *In the open-economy equilibrium, the reallocation term is $\frac{\sigma}{(\sigma-1)\theta}[(1-\theta)d\ln\lambda + \theta d\ln S]$ and is always negative.*[15]

PROOF: Appendix B.3.

This corollary presents two important features under the special case (3). First, compared to the closed economy, an open economy with any level of finite iceberg trade cost always has a lower welfare— so long as there is selection into exporting. Second, in an open economy equilibrium, a marginal reduction of iceberg cost always brings about a negative reallocation effect, so long as it results in a higher fraction of exporters in equilibrium, thus worsening misallocation with the reduction in trade costs.

The intuition for why the open economy has a lower welfare than in the closed economy is made transparent by this special case: under homogenous productivity, the efficient allocation for the closed economy is that firms have identical market shares. When the economy opens up to trade, under homogenous productivity, efficient allocation is that either all firms export or none of them export, and hence the ex-post should also be equal market shares for all firms. However, with distortions, the relatively subsidized firms produce more than in the efficient case, with the dispersion of sales (employment) reflecting the distortions. When opening up to trade, it is the relatively subsidized firms that export, in turn making the firm's distribution even more skewed—misallocation is exacerbated. The share of labor required in producing domestic goods ends up being less than the domestic output share. When firm selection is purely driven by distortions, allocative efficiency deteriorates when moving from autarky to an open economy.

Point 2 in the above Corollary focuses on a local change in trade costs in the open econ-

[15]In the welfare expression, the second and third terms cancel out, and the welfare change is $\frac{\sigma}{(\sigma-1)\theta}[-d\ln\lambda - \left(\sigma - 1 + \frac{\sigma\gamma_s}{\sigma-1}\right)d\ln\lambda + \left(\sigma - 1 + \frac{\sigma\gamma_\lambda}{\sigma-1}\right)d\ln S] = \frac{\sigma}{(\sigma-1)\theta}[-d\ln\lambda + (1-\theta)d\ln\lambda + \theta d\ln S]$, where the sign of $(1-\theta)d\ln\lambda + \theta d\ln S$ is always negative.

omy equilibrium. The difference in the change in the share of expenditure on domestic goods and the share of variable labor used in producing domestic goods constitutes a sufficient statistics for the welfare change. This change of welfare reduces to two terms: the standard ACR term, where $-d \ln \lambda > 0$, and the reallocation term, which are always negative. Overall, the change in welfare in the open economy equilibrium displays a U-shape pattern. Hence for a marginal change in trade cost, the welfare change can be negative or positive: for high levels of trade cost, there is a welfare loss; and for low levels of trade cost, there is a welfare gain. The reason is that firm selection driven by distortions are less significant when trade costs are small (at zero trade cost all firms export), and the welfare gains dominate the losses associated with reallocation. Nonetheless, when firms selection purely driven by distortions, open always have lower welfare than closed, and the reallocation term are always negative.

Having established sufficiency conditions in the special case, we can now examine a more general case. The necessary condition for the resource allocation term to be negative is: either $\gamma_s \leq \gamma_\lambda$ or $d \ln S \leq d \ln \lambda$. Intuitively, misallocation happens when the input elasticity is smaller than the revenue elasticity, i.e. more resources are used to produce the same unit of revenue. The following corollary presents a sufficient condition for $\gamma_s \leq \gamma_\lambda$ for a more general distribution for productivity and distortions.

Corollary 2. *Suppose (τ, φ) are jointly log-normal with standard deviations of σ_τ and σ_φ and correlation ρ. When $\sigma_\tau \geq \frac{\sigma-1}{\sigma}\rho\sigma_\varphi$, then;*

1. *The cumulative variable labor share distribution stochastically dominates the cumulative sales share distribution according to the likelihood ratio order.*
2. *The hazard functions $\gamma_s \leq \gamma_\lambda$ and shares $S \leq \lambda$ at any cutoff, hence moving from a closed economy to an open economy, the reallocation term is always negative.*

PROOF: Appendix B.4.

Under the condition $\sigma_\tau \geq \frac{\sigma-1}{\sigma}\rho\sigma_\varphi$, the cumulative labor share distribution stochastically dominates the cumulative sales share distribution according to the likelihood ratio order. Theoretically, we can prove that moving from a closed to open economy, the labor share used to produce exports is always greater than the export share of total sales, or in other words, the share of variable labor used in producing domestic goods is always smaller than the share of expenditure on domestic goods, i.e., $S \leq \lambda$. This implies that going from a closed to open economy $d \ln S$ is more negative than $d \ln \lambda$.

Intuitively, recall that cutoffs for production or exporting are related to firm profits, which now depend on (φ, τ), and the cumulative labor share and sales share distribution are functions of different values for $\hat{\varphi}$. Since likelihood dominance implies first-order stochastic dominance, under the above condition the cumulative labor share distribution has more mass among higher profit firms than the cumulative output share distribution. Thus, when the economy opens to trade, higher profit firms start to export, the share of labor used to produce exports would exceed the export share. We illustrate this in the numerical example in the next subsection and Figure 1.

Note that when the correlation between τ, φ is negative, $\sigma_\tau \geq \frac{\sigma-1}{\sigma}\rho\sigma_\varphi$ is always satisfied, which means moving from a closed economy to an open economy, the reallocation term is always negative. This may sound counterintuitive at first glance. When the correlation between τ, φ is negative, less productive firms are highly taxed and less likely to export. But as we discussed in this section, opening up to trade occasions more subsidized firms to expand, and so the share of labor used to produce exports would exceed the export share, making the reallocation term negative. This does not rule out, however, a direct positive technological effect that dominates, so overall the loss is smaller than the one with a positive correlation. Again we illustrate this in the numerical example in the next subsection (Figure 6 (a)).

Proposition 3 applies to both symmetric and asymmetric countries and takes into consideration the impact of the Foreign distribution of firms on the Home country. It also shows the effect of domestic distortions on a Foreign country. In the case where the Foreign economy does not have distortions, the third and fourth terms in Foreign's welfare formula go to zero. The Foreign welfare is provided in the following proposition:

Proposition 4. *(**Foreign Welfare**) In the case that the Foreign economy is devoid of distortions,*

$$d \ln W_f = \frac{1}{\gamma_f + \sigma - 1}[-d \ln \lambda_f + d \ln M_{ef}].$$

Thus, Home's domestic distortions affect Foreign only through Foreign's λ_f, M_{ef}, the cutoffs φ_f^*, and hence γ_f.

Discussion. Proposition 3 is useful for making transparent the key mechanism that underlies how trade and misallocation affect welfare. It also easily relates to the existing literature, such as the ACR and MR formulations. Thus, it is the main decomposition we emphasize. However, there are various guises under which the welfare decomposition can assume. For example, using the equilibrium conditions, we can decompose welfare changes into domestic sales share, entry, and aggregate wedge $\overline{MRPL}$, i.e.

$$d\ln W = \frac{1}{\gamma_\lambda + \sigma - 1}[-d\ln\lambda + d\ln M_e] + \left(\frac{\gamma_\lambda/(\sigma-1)}{\gamma_\lambda + \sigma - 1} + 1\right) d\ln\overline{MRPL}, \tag{17}$$

where the aggregate wedge is defined as the marginal revenue product of labor, $\overline{MRPL} = PQ/L$. With a constant labor, the change in the average wedge is the same as the change of aggregate expenditure PQ, i.e. $d\ln\overline{MRPL} = d\ln PQ$. Trade changes the aggregate expenditure (and thus selection) and hence welfare in the economy. Equivalently, the welfare change can be written in the following way,

$$d\ln W = \frac{1}{\sigma-1}\left[-d\ln\lambda + d\ln M_e - \gamma_\lambda d\ln\hat{\varphi}^*\right] + d\ln\overline{MRPL} \tag{18}$$

$$= \frac{1}{\sigma-1}\left[-d\ln\lambda\right] - d\ln P_d + d\ln\overline{MRPL}, \tag{19}$$

where P_d is the price index for domestic goods. We can label the first term as gains from import, and the second term relates to loss from exit as in Hsieh, Li, Ossa, and Yang (2016). Without distortion, $d\ln\overline{MRPL} = 0$. In our model, distortions affect domestic sales share, entry, selection and resource allocation. The additional effect of $d\ln\overline{MRPL}$ and γ_λ reflect how selection in foreign market and domestic market change the joint distribution of wedge and sales, i.e. the allocation of resources.[16]

The change in the average wedge $d\ln\overline{MRPL}$ maps onto the reallocation effect in Propo-

[16]Without distortions, firms exit if and only if welfare Q improves. The reason is that real wage w/P increases as welfare improves, $Q = (w/P)L$, and domestic selection depends only on w/P. However, with distortions, the link between firm exit and rising welfare is broken. Competition arising from trade may drive out firms from either domestic or foreign market but still reduce welfare. A good example is the point 3 after the Proposition 3.

sition 3. It links to the gap between $d \ln S$ and $d \ln \lambda$. The aggregate wedge is a harmonic average of firm distortions with sales as weights. It also relates to the tariff/wedge and labor share as in Baqaee and Farhi (2019) and Edmond, Midrigan, and Xu (2018). A fall in aggregate wedge implies welfare losses—as resources are reallocated towards more subsidized firms.

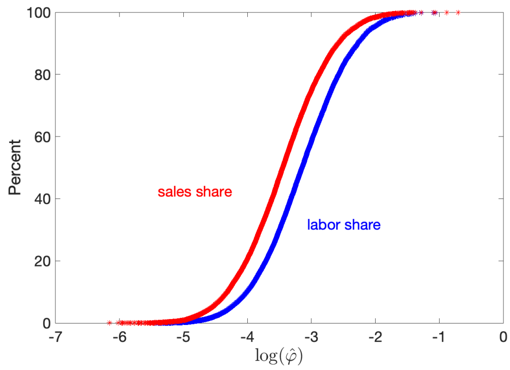
In general, all of these decompositions, (16), (17), and (19), demonstrate that for the ex-post gains from trade, one needs firm-level data to examine the change in hazard rates γ_λ and/or γ_s, the response of firm entry and changes in share gap or aggregate wedge. As discussed in MR, measuring the response of firm entry to changes in trade cost is equally challenging to estimating a partial trade elasticity and recovering the change in trade induced by a change in trade costs alone. Moreover, distortions pose an even more daunting task for measuring the welfare gain directly using firm-level data. Hazard rates γ_λ and γ_s are variant and depend on the underlying joint distribution of wedge and productivity and micro structure. Due to measurement issues and endogenous selection, we cannot directly measure the joint distribution from the data. A detailed discussion is provided in Section 3.1. Furthermore, ex-ante welfare evaluation depends on the underlying joint distribution of wedge and productivity and across different values for trade costs in the model as they change the reallocation effect. We therefore will use our model to estimate the underlying parameters and the corresponding gain or loss from trade for both ex-ante and ex-post exercises in Section 3.

2.3 Numerical Example

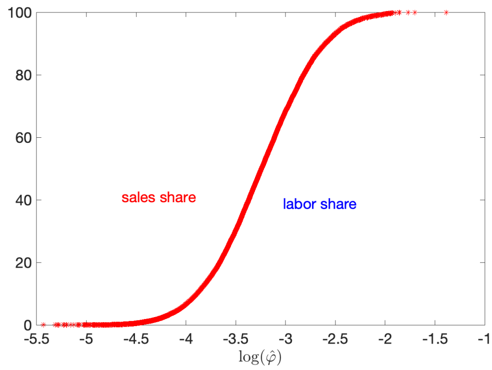
To unpack the theoretical results and to provide more intuition for the mechanisms that underpin these results, we next turn to a numerical example of the benchmark model with symmetric countries. The joint distribution between productivity and distortions is taken to be joint log-normal with standard deviations of $\sigma_\tau = \sigma_\varphi = 0.5$ and correlation of φ and τ of $\rho = 0.8$. The elasticity of substitution $\sigma = 2$, and the fixed costs are $f = 0.03$, $f_x = 0.035$, $f_e = 0.01$.

Corollary 2 applies here as the distribution of (φ, τ) and the parameters satisfies its conditions. We plot the cumulative variable input and sales share under any $\log(\hat{\varphi})$ in panel

Figure 1: Accumulated Labor Share vs Sales Share in a Market



(a) With Distortion



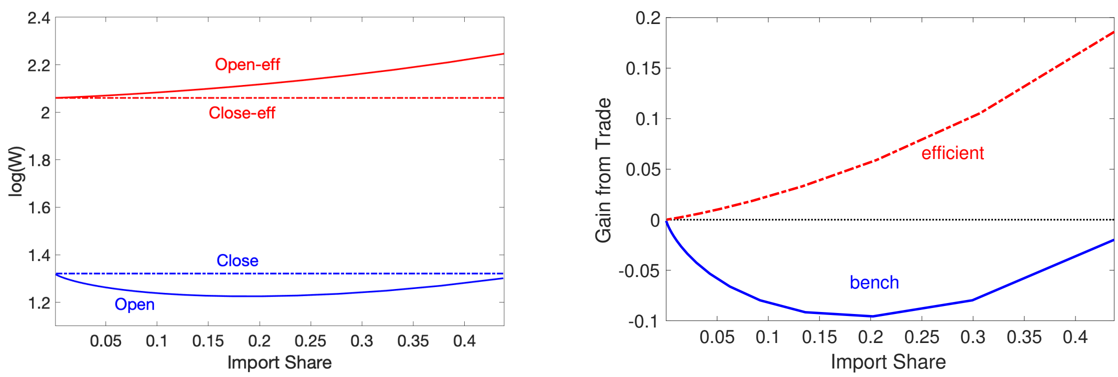
(b) Without Distortion

(a) of Figure 1. According to Corollary 2, the cumulative variable input share distribution stochastically dominates the cumulative sales share distribution according to the likelihood ratio order, which implies first-order stochastic dominance. In contrast, without distortions with $\tau = 1$, these two distributions are identical as shown in panel (b) of Figure 1. When the economy opens to trade, firms that export are those with high profit and also use a large share of labor to produce. Overall, the share of labor used to produce exports would exceed the export share, worsening the misallocation of resources.

The example helps illustrate a few points. First, welfare (Eq. 15) can fall when the economy opens up to trade. Figure 2 (a) plots the level of welfare against import shares under the alternative scenarios: the efficient case without distortions, the case with distortions, and when the economy is closed or open. Three observations immediately follow: 1) that there is a welfare loss in the case with distortions compared to the case without; 2) opening up to trade leads to welfare gains in the efficient case; however, 3) opening up engenders a welfare *loss* in the presence of distortions. Taking the differences between the open and close economy in either case, with or without distortion, we plot the welfare change after trade in Figure 2 (b). It is clear that there is welfare loss with distortions in our benchmark.

Second, we plot the welfare decompositions according to the welfare formula (16) in Figure 3 (a), which displays the three components in our benchmark model— the ACR term, entry M_e, and reallocation. Through the impact of technology, welfare increases with lower trade cost, i.e the ACR term increases with the import share. Trade, however, exacerbates the misallocation and reduces allocative efficiency. Hence, the reallocation term becomes

Figure 2: Welfare and the Change from Trade



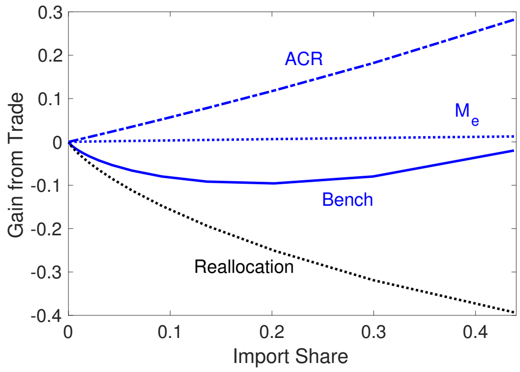
(a) Welfare (b) Gain from Trade

more and more negative as the import share increases. Overall, the resource misallocation effect dominates, and the benchmark shows welfare losses with trade. In the current example, the entry effect (the change in varieties) plays a very small role. It is negative in the case without distortions but practically unchanged in the case with distortions. However, the channel we emphasize 'reallocation' has a first-order impact in the case with distortions. In the case without distortions, there is no misallocation and the resource allocation term remains zero. The entry effect becomes more negative with trade, which mitigates the increase in ACR. Overall, there is a welfare gain after trade.

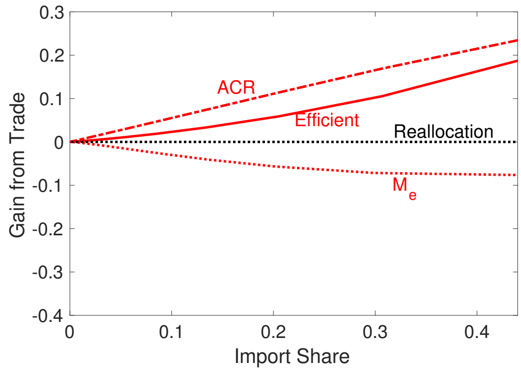
Third, the numerical example also demonstrates that using import shares to infer welfare changes can gives rise to markedly different results when there are distortions. Figure 3 (a) shows that ACR invariably predicts gains to trade, rather than losses. Figure 3 (b) shows that in the absence of distortions, ACR is a good approximation for welfare in the efficient case. But using ACR under distortions leads to a large departure: under these benchmark results there are welfare *losses* rather than gains. Using aggregate observables to infer welfare gains as in ACR can thus be very misleading in the presence of distortions.

Note that in this numerical example, the two countries are symmetric, and both face domestic distortions. The assumption of symmetry abstracts from terms of trade effect and highlights the role of misallocation in generating loss from trade. Specifically, Home suffers a loss from trade is not because Home is subsidizing firms' exports and Foreign gains due to a terms of trade effect. This symmetric example emphasizes that loss from trade comes from the deterioration of resource allocations.

Figure 3: Welfare Decomposition



(a) With distortion

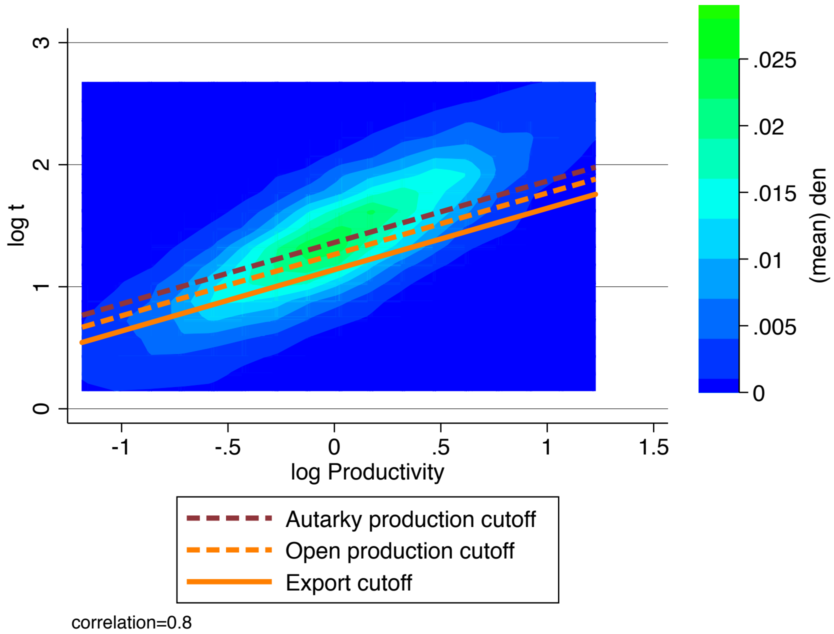


(b) Without distortion

Fourth, we can examine the selection mechanism at the micro level, which explains how trade can reduce aggregate efficiency. In the same numerical example, Figure 4 illustrates how distortions affect firm selection. The density of firms is shown by a heat map of firms that lie along a positively sloped distortion-productivity line. It is clear that the productivity cutoff for production and exports is no longer determined solely by productivity, but also by domestic distortion. Only firms below the cutoff line can operate. In this figure, a large mass of highly-productive firms are excluded from servicing the market altogether. As the economy opens up, the cutoff line is shifted further downward. Even if firms have the same level of productivity, some with higher taxes may be displaced while those with lower ones will survive. This downward shift of the cutoffs allows for some low productivity and high subsidy firms to survive and gain market share.

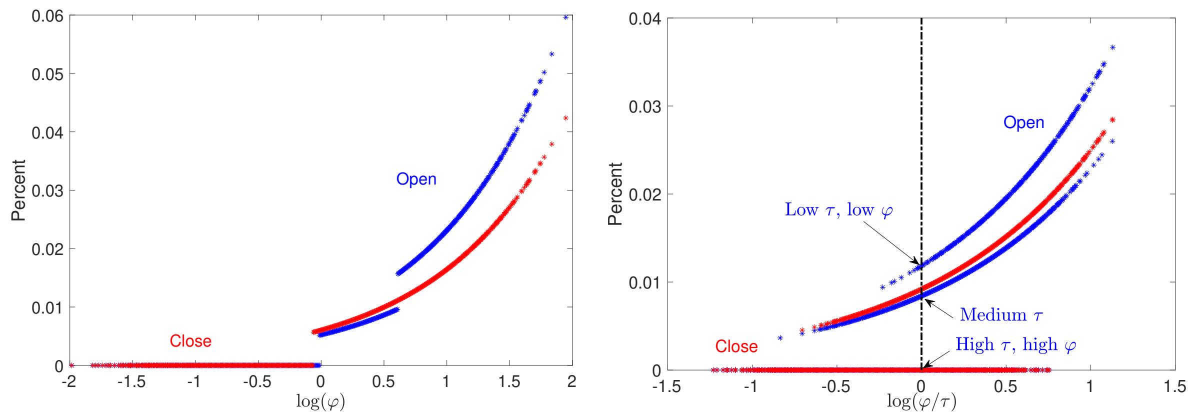
Another way to show the impact on selection is to examine firms' market share. The two panels in Figure 5 plot the market share of firms, both in the closed and open economy. The left panel is the case without distortions. Without distortion, the marginal cost is the inverse of the productivity φ. Firms with the same productivity level have the same marginal cost; their market share, above a cutoff productivity, rises with their productivity. Comparing the blue and red lines show that above the export cutoff, more productive firms have higher market shares in the open economy than in the closed economy, demonstrating that these firms expand under trade liberalization. This happens at the cost of displacing other less productive firms' market share, or driving them out of the market entirely. Here, the example clearly demonstrates that resources move from less productive to more productive

Figure 4: Cutoffs



firms as an economy opens up to trade.

Figure 5: Selection Effects



(a) Without distortion

(b) With distortion

The right panel shows the firm's market share in the case with distortions. Firms may share the same marginal cost τ/φ and face the same potential revenues. However, their after-tax profits may differ, and thus their market share can also differ. Consider the point at which $\log(\varphi/\tau)$ is at 0. At this point, a firm with high, medium and low level of productivity face the same marginal costs. However, the high productivity firm is also subject to high taxes and thus low after-tax profit, and does not make the cut for production.

The medium-tax-medium-productivity firm has positive market share but loses out to the low-tax-low-productivity firm when the economy opens up. Resources are reallocated from the more productive to the less productive firms. Also, there is no longer a neat line up of market shares according to productivity: there is a wide range of productivities for which production is excluded.[17] Aggregate welfare effect depends on how trade alters the aggregate domestic labor share and sales share.

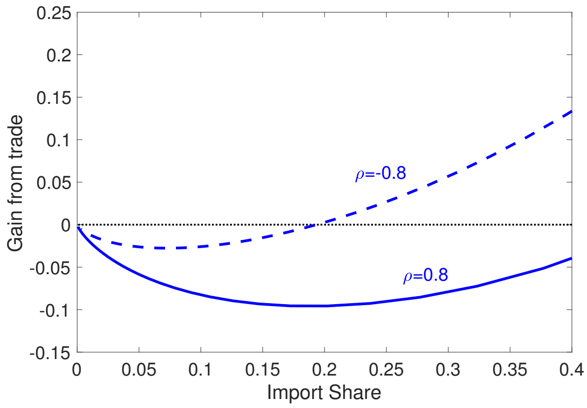
Distribution of Distortions. The distribution of distortions is an important determinant to the gains to trade. There are two key parameters: ρ, the correlation of τ and φ, and σ_τ, the dispersion of τ. Figure 6 (b) compares the gains from trade under different σ_τ, while the other parameters remain the same as in the benchmark example. The welfare gain (loss) from trade is always larger (smaller) when σ_τ is smaller.

The correlation of distortion and productivity is important insofar as a higher correlation means that more productive firms are more likely to be excluded from the market. But reductions in welfare is possible even when the correlation is negative. The reason is that for any given productivity, it is always the more subsidized firms that can export, and the highly taxed ones that exit— leading to a possible worsen of misallocation. In fact, as shown in Corollary 2, when the correlation is negative, more productive firms are highly subsidized. Exporters are those more productive and highly subsidized ones, hence their labor share are larger than sales share, and the reallocation term is always negative. Overall effects combine the positive "technology" effect and the negative reallocation effect. Figure 6 (a) illustrates this. It compares the gains from trade for $\rho = 0.8$, under our benchmark numerical example, and for $\rho = -0.8$, where productivity and distortion are highly negatively correlated. Under $\rho = -0.8$, the welfare gain (loss) from trade is always larger (smaller) than that in the case of $\rho = 0.8$. But when the import share is below 20%, there are still losses from trade even under a negative correlation.

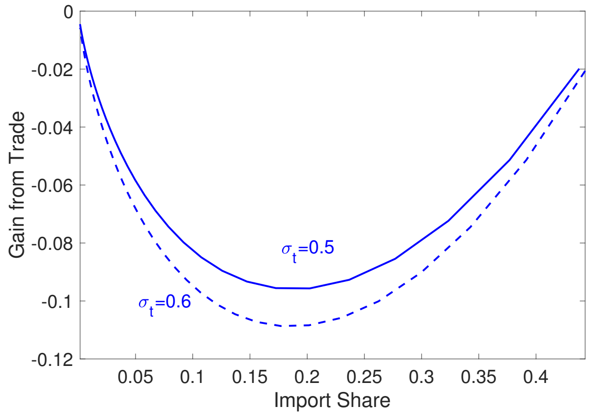
In sum, the size of welfare loss after opening up depends on the correlation of φ and τ and the dispersion of τ. The firm level data helps us identify these parameters. Specifically, in the quantitative section, we will use the firm-level output and use its dispersion and its

[17] This is also true if the distortions are input wedge on all the labor a firm uses. Firms face higher input wedge would have a lower profit in a market.

Figure 6: Gains/Loss from Trade



(a) Correlation Matters



(b) Dispersion Matters

correlation with firm inputs to estimate ρ, σ_τ, and σ_φ.

3 Quantitative Analysis

This section presents estimates of the quantitative effects of trade liberalization when including domestic distortions. The two countries, Home and Foreign, are calibrated to data corresponding to China and the U.S.. The U.S. is taken to be the relatively undistorted economy. In what follows, we first demonstrate why distortions and productivity *cannot* be measured directly from the data—a customary approach adopted in past and present works. Instead, our approach is to take the moments related to distortion and productivity in the data, use them to parameterize our model, and then estimate the welfare gain or losses from trade liberalization. We also conduct out-of-sample tests by examining the differential patterns among exporters and non-exporters in both the cross-section and the time series dimension and compare the results to key implications of our model. Lastly, we consider extensions to our benchmark model. Our aim is not to provide a full-fledged quantitative account of China's trade liberalization experience. For that, one would need a much richer model with a complex set of mechanisms. Instead, the main purpose is to use China as an example to demonstrate the large quantitative and potentially qualitative differences that may arise under a model with distortions, compared to the standard model without distortions. After all, the presence of distortions is a prominent feature of not only the Chinese economy, but also of many other developing countries for which trade liberalisation is most vigorously studied.

3.1 Data and Measurement

The data for Chinese firms comes from an annual survey of manufacturing enterprises collected by the Chinese National Bureau of Statistics. The dataset includes non-state firms with sales over 5 million RMB (about 600,000 US dollars) and all of the state firms for the 1998-2007 period. Information is derived from the balance sheet, profit and loss statements, and cash flow statements, which incorporate more than 100 financial variables. The raw data consist of over 125,858 firms in 1998 and 306,298 firms by 2007.

Our strategy is to use the observed distributions of inputs, outputs, and export participations from Chinese firm data to estimate the joint distribution of distortions and productivity in conjunction with other parameters in the model. We do not recover the joint distribution directly from the data for two reasons. First, neither a firm's productivity nor its distortions can be measured directly. Second, firm selection affects the observed joint distribution of wedge and productivity. The observed joint distribution in the data is the ex-post one after selection, rather than the underlying one. The importance of these two issues merits a full elaboration below. The alternative strategy that we choose to adopt is then detailed in Section 3.2.

The customary way to recover a firm's productivity or distortion is to use its value-added per input. This measurement, however, could be contaminated with the presence of distortions or fixed cost of producing. Using the first order condition of a firm's optimization problem (1), we can write the value added per input as

$$\frac{pq}{\ell} \propto \tau \left[1 - \frac{f}{\ell(\varphi, \tau)} \right], \tag{20}$$

where f also includes exporting fixed cost if firm exports. If there are no wedges, $\tau = 1$, the value added per input increases with input ℓ, which in turn increases with a firm's physical productivity, as in Melitz. If there are no fixed costs, $f = 0$, the value added per input actually measure the firm's wedges, as in HK. With both wedges and fixed costs, the value added per input not only depends on the productivity of the firm but also on the true wedge τ. For this reason, we cannot use the value- added per input to measure the firm's productivity or its wedges.

The second reason for which one cannot recover the joint distribution directly from the data is firm selection. The observed dispersion and correlation of some measured wedge and productivity pertains to only operating firms. Hence, it is confounded with an endogenous selection mechanism. For instance, even if the underlying correlation were negative, the selection mechanism can induce the observed correlation to *become positive*, for the simple reason that high-taxed firms must be more productive in order to stay in the market. The selection mechanism will strengthen any underlying correlation between the two variables. For the same reason, the observed dispersions of the two variables are also the ones after selection has taken place. In order to compute the impact of distortions on welfare and productivity gains, one would need to know the underlying correlation and dispersion, and therefore one would need micro data and a structural model to uncover it. This is exactly the approach we adopt, detailed in the following subsection.[18]

3.2 Parameterization

Table 1 reports the parameter values. We set the elasticity of substitution between varieties σ to be 3, the one taken in HK. This value is consistent with the estimates from plant-level US manufacturing data in Bernard, Eaton, Jensen, and Kortum (2003). The Home labor L is normalized to 1. Given that Foreign affects Home only though aggregate variables, we can assume that Foreign is absent of distortions, while taking $f_e, f, f_x, \tau_x, \sigma_\varphi$ to be the same as those in Home.

The remaining 9 parameters are estimated jointly, to match the model moments with their data counterparts. Table 1 reports the estimated parameters and the moments in the data and model. The moments we choose are the ones that are most relevant and sensitive to variations in model parameters. Clearly, every parameter matters for the general equilibrium and affects other moments. However, there is by and large a clear correspondance between certain parameters and moments. The parameter most relevant for matching the

[18]Note that our approach is different from the existing literature, for example Costa-Scottini (2018) and Ho (2010), which assume a perfect correlation of (log) productivity and (log) wedges and take the joint distribution of average revenue product of labor and measured productivity directly from the data.

Table 1: Parametrization and Moments

Panel A: Parameters		Panel B: Moments		
Endogenously chosen	Value	*Targeted moments*	Data	Model
Entry cost f_e	0.2	Fraction of firms producing	0.85	0.85
Fixed cost of producing f	0.015	Mean-lowest 5% $\ln(k^{\alpha}\ell^{1-\alpha})$	1.82	1.53
Fixed cost of export f_x	0.12	Fraction of firms exporting	0.30	0.28
Iceberg trade cost τ_x	1.5	Export intensity	0.41	0.42
Std. productivity σ_{φ}	1.2	Std. $\ln VA$	1.20	1.26
Std. distortion σ_{τ}	0.9	Std. $\ln(VA/(k^{\alpha}\ell^{1-\alpha}))$	0.93	0.84
Corr(distortion, productivity) ρ	0.86	Corr$(\ln VA, \ln(VA/(k^{\alpha}\ell^{1-\alpha})))$	0.41	0.35
Mean foreign prod $\mu_{f\varphi}$	5.5	Relative GDP of U.S. to China	1.79	1.77
Foreign labor L_f	0.2	Relative labor of US to China	0.20	0.20
Assigned parameters				
Elasticity of substitution σ	3			
Home labor L	1			

Note: VA denotes value added, k capital, ℓ employment, *Std* standard deviation, Corr correlations.

fraction of surviving firms is the entry cost f_e, as $\omega_e E[\pi(\varphi,\tau)] = wf_e$. Lower entry costs induces more entrants to pay the costs, and the result is a lower fraction of survivors. Next, to identify the fixed cost f, one needs only to turn to the smallest firms, which have their profit just about cover fixed cost. That is, the after-tax profit $\pi = wf$ and $w\ell_{min} = (\sigma-1)wf + wf$, and the mean of firms' labor $w\ell_{mean} = (\sigma-1)w(f_e/\omega_e + f) + wf$. Hence, the difference between mean and lowest 5% of input helps identify f.

We choose foreign labor L_f to 0.2 to match the relative labor force of US to China. We calibrate f_x and τ_x to match the export participation and intensity in Chinese manufacturing. The resulting parameter $\tau_x = 1.5$ is inline with the estimate of 1.7 in Anderson and Van Wincoop (2004), and the 1.83 in Melitz and Redding (2015). The dispersions in productivity and distortions, and their correlation are important for matching the observed joint distribution between value-added and inputs in the data.

In the model, labor represents inputs. However, the corresponding part in the data is the composite inputs $k^{\alpha}\ell^{1-\alpha}$, aggregated under a Cobb-Douglas function. This assumption works both for the variable input and the fixed cost in the model. For a firm in industry j, its output q is produced with variable labor ℓ_v and capital k_v using the form $q_j = \varphi k_v^{\alpha_j}\ell_v^{1-\alpha_j}$

with an industry specify labor share $1-\alpha_j$. These industry labor shares come from the U.S. NBER productivity database, which is based on the Census and the Annual Survey of Manufactures (ASM). The reason that labor shares are not computed from Chinese data is that the prevalence of distortions would affect these elasticities, and industry-level elasticities and distortions cannot be separately identified. Different from HK, we take a firm's total employment to measure ℓ_j rather than the firm's wage bill. This addresses the problem that Chinese wage data implies too low of a labor share as measured by input-output tables and the national accounts. We define the capital stock as the book value of fixed capital net of depreciation.

Table 1 shows that the discrepancy between our model and data moments is reasonably small, though we underestimate the dispersion in distortions and slightly overestimate the dispersion in size. An important variable is the correlation between value added and $VA/(k^{\alpha}\ell^{1-\alpha})$, $Corr(\ln VA,\ \ln VA/(k^{\alpha}\ell^{1-\alpha}))$. This variable is more positive the higher is $\rho\frac{\sigma_{\varphi}}{\sigma_{\tau}}$, where ρ is the underlying correlation between distortions and productivity. A higher underlying correlation and a lower dispersion in distortions raise the observed correlation between value added and inputs. Lastly, we choose the mean of foreign productivity $\mu_{f\varphi}$ to match the relative GDP of US to China.

3.3 Implied Gains from Trade and Loss in TFP

Table 2 reports the gains from trade and efficiency losses for both Home and Foreign. The upper panel compares welfare and TFP in the open economy to those in the closed economy. In the benchmark estimation, the gains from trade for Home is 4.4%. Without distortions, the gains from trade is more than doubled (9.8%). Foreign's gain from trade is about 8.2% when Home has domestic distortions. Eliminating Home distortions allows Foreign to benefit more—a 19% of welfare gain.

Note that the general formulation of welfare given in Proposition 3 holds for asymmetric countries. Thus, one can decompose the welfare change according to its main equation, (Eq.16). In the benchmark case, the import share is 30.8%, which implies that the change in domestic output share is $-d\ln\lambda = 0.368$, and the ACR term in the equation is 17.9%. But there is a large and negative reallocation term showing up in China, amounting to

Table 2: Welfare and TFP

	Open relative to close			Decomposition	
	Welfare	TFP	Import Share	ACR	Reallocation
Home (%)					
Benchmark	4.4	-2.9	30.8	17.9	-18.2
No-distortion	9.8	13.3	20.8	11.4	0
Foreign(%)					
Benchmark	8.2	12.9	17.9	9.0	0
No-distortion	18.9	13.3	35	19.1	0
	TFP loss: Distortion relative to no-distortion				
	Overall loss	Misallocation	Entry-selection		
Benchmark	140.4	119.2	21.2		
Home Closed-Economy	124.2	118.7	5.4		

−18.2%. Taking only the ACR component would overestimate the gains to trade by 407%, according to our model.[19]. Moreover, if one were to ignore the presence of distortions and followed the usual approach to ACR using a trade elasticity– for instance of 4, estimated in Simonovska and Waugh (2014), then the resulting welfare gains would be 9.2% – more than double the gains that our model predicts. As Melitz and Redding (2015) show, trade elasticities vary with trade costs, so whether a partial, a full or an average theoretical trade elasticity is used will matter for the size of the ACR gain. But regardless, the ACR term is positive no matter which elasticity is utilized. What we show here is that our 'reallocation' channel leads to a sizeable welfare loss.

We assume here that Foreign does not face any distortions, and thus the reallocation term for Foreign is zero according to our Proposition 3. Since Foreign has a smaller import share than Home, one would arrive at the conclusion that more gains would accrue to Home than to Foreign, according to ACR. But our benchmark model predicts the opposite: Foreign has gains to trade that is almost double that of Home, whereas using conventional ACR, Foreign's gain is only half of that of Home. Without distortions, the gains to trade would double for both economies, suggesting that countries would gain more from trade by undertaking domestic reforms.

We also compare TFP before and after trade liberalization. To compute TFP, we need

[19]Note that in the decomposition in Proposition 3, the ACR component is related to the γ_s, the elasticity of the cumulative variable labor share within the domestic market for firms above the domestic cutoff, with respect to the cutoff. It is much smaller than γ_λ, the elasticity of the cumulative sales share within the domestic market for firms above the domestic cutoff, with respect to the cutoff. See corollary 3.

to define real GDP. As shown in Burstein and Cravino (2015), there are different ways to construct price index and therefore real GDP. Here, we define the aggregate producer price as $PPI = (\int p_i^{1-\sigma} di)^{\frac{1}{1-\sigma}}$, where p_i is the price charged by producer i. Our benchmark results show that opening up leads to a 3% TFP *loss*. In contrast, without distortions, TFP increases by 13.3%. Hence, contrary to the standard predictions, trade liberalization can exacerbate rather than improve resource allocation, causing a decline rather than a rise in TFP. As Foreign has no distortions, its TFP levels are basically the same between the two models.

The lower panel of Table 2 reports Home's TFP losses due to distortions, both for a closed and open-economy case scenario. TFP loss is defined as the difference between TFP under the case of distortions and no distortions. Different from Hsieh and Klenow (2009), our efficient TFP considers endogenous entry. We are therefore able to decompose the deviation of TFP from its efficient level into a misallocation effect among a fixed set of operating firms, and a misallocation effect generated by entry and selection into producing and export:

$$\log TFP_{eff} - \log TFP = \underbrace{\log TFP_{FX} - \log TFP}_{\text{misallocation loss}} + \underbrace{\log TFP_{eff} - \log TFP_{FX}}_{\text{entry and selection loss}},$$

where TFP_{eff} pertains to the case without distortions, TFP_{FX} corresponds to the level in the case without distortions among producing firms but where M, φ^*, and φ_x^* are fixed as in a distorted economy.[20]

Not surprisingly, there are large TFP losses for Home with domestic distortions than

[20] Here are the definitions of TFP_{eff} and TFP_{FX} in the closed economy,

$$TFP_{eff} = \frac{\sigma-1}{\sigma}\left[M^{eff}\int_{\varphi^{eff*}}^{\infty}\varphi^{\sigma-1}\mu^{eff}(\varphi)\,d\varphi\right]^{\frac{1}{\sigma-1}}, TFP_{FX} = \frac{\sigma-1}{\sigma}\left[M\int_{\varphi^*}^{\infty}\varphi^{\sigma-1}\mu(\varphi)\,d\varphi\right]^{\frac{1}{\sigma-1}}.$$

and in the open economy,

$$TFP_{eff} = \frac{\sigma-1}{\sigma}\left[M^{eff}\int_{\varphi^{eff*}}^{\infty}\varphi^{\sigma-1}\mu^{eff}(\varphi)\,d\varphi + M^{eff}\int_{\varphi_x^{eff*}}^{\infty}(\frac{\varphi}{\tau_x})^{\sigma-1}\mu^{eff}(\varphi)\,d\varphi\right]^{\frac{1}{\sigma-1}}$$

$$TFP_{FX} = \frac{\sigma-1}{\sigma}\left[M\int_{\varphi^*}^{\infty}\varphi^{\sigma-1}\mu(\varphi)\,d\varphi + M\int_{\varphi_x^*}^{\infty}(\frac{\varphi}{\tau_x})^{\sigma-1}\mu(\varphi)\,d\varphi\right]^{\frac{1}{\sigma-1}}$$

without. In the closed-economy case, eliminating these distortions would increase China's TFP by 124%. The TFP losses due to the presence of distortions are larger, 140%, in the open economy benchmark model where China has an import share of more than 30%. Looking at its decomposition, the majority of the losses appears to come from misallocation among existing firms. One reason is that the share of surviving firms is high according to available data, but the survival rate could be overestimated as the dataset includes firms only of a certain scale. If smaller firms were observed, then it is possible that a lower survival rate would make the losses coming from the entry and selection margin larger.

3.4 Selection through Export: Out-of-Sample Tests and Extensions

To provide a further check on the mechanisms that are highlighted, we proceed to undertake a series of out-of-sample tests and extensions. In particular, one can compare the differential patterns among exporters and non-exporters in both the cross-section and the time series dimension in the data and in the model. In order to do so, we construct the customary 'TFPR' and 'TFPQ' measures following HK, in both the data and the out-of-sample model. Note that they *relate* to distortions and productivity, but are not the underlying distortions and productivity in the model, as we have explained previously.

We construct the TFPR with the weighted average of average revenue product of labor and average revenue product of capital,

$$\log(TFPR_{ji}) = \alpha_j \log(ARPL_{ji}) + (1-\alpha_j)\log(ARPK_{ji}).$$

Specifically, the relative average revenue product of labor ($ARPL_{ji}$), is calculated as $\log\left(p_{ji}q_{ij}/\ell_{ij}\right) - \log\left(\overline{ARPL_j}\right)$ where $\overline{ARPL_j}$ is the industry mean of the average product. Similarly, we construct relative average revenue product of capital. In our model with fixed cost, the average revenue product is a biased measurement of marginal revenue of product. Hence TFPR is affected by the wedge, but also affected by the productivity of a firm, as shown in equation (20). We find large dispersions in TFPR in China, similar to the levels in HK for the year 1998 and 2007. Measured TFPR have come down over time, between 1998 and 2007, as evident in Table A-1.

The TFPQ is related to the physical productivity of a firm and is measured with

$$TFPQ_{ji} = \left(P_j^{\sigma-1} Q_j\right)^{\frac{1}{1-\sigma}} \frac{(p_{ji} q_{ji})^{\frac{\sigma}{\sigma-1}}}{k_{ji}^{\alpha_j} \ell_{ji}^{1-\alpha_j}}, \tag{21}$$

with its deviation from the industry mean. In our model, the measured TFPQ uses the similar formula but replacing $k_{ji}^{\alpha_j} \ell_{ji}^{1-\alpha_j}$ with total input used in a firm. Again both in the data and the model, we use total inputs instead of variable inputs, *hence TFPQ is not φ in our model, but is affected jointly by firm productivity, distortion and the selection into foreign market*.

The first two columns of Table 3 reports the data and the model regressions of measured TFPR on measured TFPQ and a dummy of exporters. Both the model and the data exhibit a pattern whereby exporters face a lower TFPR. Note that the differences between exporters and non-exporters were not targeted. There is a stronger selection effect in the model than in the data—exporters' marginal product is about 64% lower than non-exporters in the model, compared to 26% in the data. [21]

Exporters may face different distortions, which is absent from our benchmark model. To examine whether Chinese firm characteristics change when they become exporters, we examine the relationship between measured TFPR and firm export status in both cross section and the time series dimension.

We use the time series data to check whether distortions change when firms enter the export market. This will help us understand whether differences between exporters and non-exporters stem from selection or additional/different distortions when exporting. Throughout the sample period, we sort exporters into three types: 'always exporters' are those who are exporting throughout the sample years 1998 to 2007, 'starters' are those who started to export after 1998, and 'stoppers', who stop exporting sometime in the interim years. Entry effect measures the percentage difference of TFPR for starters, between the post- and pre-exporting entry periods. Exit effect measures percentage difference of TFPR

[21] Table A-2 investigates what factors are systematically related to TFPR for interested readers, and reports the regression results of the TFPR of a firm on a set of variables beyond TFPQ and exporting status. The coefficient on TFPQ is large and significant; 1 percent increase in relative *TFPQ* is associated with a 0.7 percent increase in relative TFPR. Moreover, more than half of the variation is explained by TFPQ alone. Given TFPQ, exporters have lower TFPR on average.

Table 3: TFPR, TFPQ, and exporting status, data vs model

	(1)	(2)	(3)	(4)	(5)	(6)
VARIABLES	Data	Benchmark	Data	Benchmark	Export rebate	Different τ in export
entry effect			-0.104***	-0.050	-0.103	-0.09
			(-12.69)			
exit effect			0.0315***			
			(4.574)			
starter			-0.101***	-0.429	-0.400	-0.08
			(-21.74)			
stopper			-0.0891***			
			(-20.98)			
always exporters			-0.301***	-0.768	-0.791	-0.324
			(-23.47)			
log(TFPQ)	0.636***	0.652	0.638***	0.653	0.654	0.613
	(250.5)		(254.9)			
exporters	-0.264***	-0.637				
	(-24.14)					
Constant	-3.258***	0.401	-3.255***	0.414	0.425	0.714
	(-106.2)		(-107.0)			
Observations	1,587,629		1,584,242			
R-squared	0.823		0.826			
Time FE	Yes		Yes			
Industry FE	Yes		Yes			

Note: Robust t-statistics in parentheses, *** p<0.01, ** p<0.05, * p<0.1
The dependent variable is measured average product, which is $\log(TFPR)$ in the data and $\log(VA/(\ell_v + f))$ in the model.

for stoppers, between the post- and pre-exporting exit periods. As shown in Column 3 of Table 3, in the data, 'always exporters' have a lower TFPR. Firms' measured TFPR decrease when they start exporting and increase when they exit.

Notice that in our benchmark model that although the firm-level distortions are exogenous, the measured TFPR do change after firms start exporting— due to the fixed costs of exporting. Column 4 shows this mechanism. In our benchmark model, if trade costs decreases (from their 1998 level to their 2005 level in Section 3.5), some firms would start to export, and their measured TFPR would decrease— as shown in the entry-effect of column 4. Both 'always exporters' and 'starters' have lower measured TFPR, which are consistent with the data. Our benchmark model shows a relatively large selection effect since the model does not have as large heterogeneity as in the data, and a relative small entry effect. In reality, the change in a firm's TFPR after exporting could be driven by multiple factors. For example, exporters face different distortions from non-exporters, or there are endogenous distortions that change with trade liberalization. We proceed to explore these possibilities. Here we consider model extensions with export rebate and allow for different distortions when exporting. In section 4.1 we consider an example of endogenous distortion.

In Column 5, we introduce a 10% tax rebate after exporting. This generates further decrease of the measured TFPR upon exporting and an even larger reallocation effect. In this case, export subsidies from tax rebate benefit the Foreign country, and Home welfare gains from trade decreases to 1.6% and TFP loss increases to 4.5%.

Column 6 considers an extension that firms face different distribution of distortion when exporting. We use the standard deviation of export intensity to discipline the dispersion of the additional wedges when exporting. To be more precise, we estimate the mean and standard deviation of the extra export distortions to match the standard deviation of export intensity and the regression coefficient on always exporters. When the trade cost is reduced as in column 4, the non-targeted coefficient on starter and the entry effect are similar to the data. Home welfare gains from trade further decreases to 1.1% in this extension. The reason is that there are additional, direct distortions on firm selection into Foreign market, when opening up to trade. Overall, if there are different distortions associated with exporting, then the quantitative exercises show that losses are larger due to larger dispersion of wedges, larger selection and reallocation effects.

3.5 Decomposing China's Growth from 1998-2005

The rapid growth in China over the last four decades has been one of the most remarkable phenomena the world has witnessed in recent history. In between 1998 and 2005, its real GDP increased by 57%. Accompanying this development was a combination of domestic reforms and opening up programs—policies that fostered trade and FDI inflows. As a result, both trade and technological progress increased over time, while domestic distortions concurrently fell. A natural question is how much of the growth is attributed to trade over this period. Other competing factors include technological improvement, factor accumulation, and domestic reforms—that is, the allocative gains associated with a reduction in distortions. In what follows, we perform a quantitative analysis to answer this question. Specifically, we recalibrate the model parameters for the year 1998 and compare the implied GDP and TFP levels to those in the benchmark year, 2005. Overall, our results attribute the majority of China's GDP and TFP growth to technological improvement, capital accumulation, and a mitigation of distortions. Trade alone contributes to only about 8% of GDP growth.

Table 4 reports the moments for both 1998 and 2005. The starting year is taken to be 1998, as it is the first year in which firm-level data is available, and three years before China joined the WTO. Compared to the year 2005, trade intensity was significantly lower in 1998, both in terms of the fraction of firms that export, and also the export intensity of these firms. Distortions were large in the earlier years, as partly seen by the fact that the dispersion of TFPR was about 20% higher in 1998 compared to 2005. These imply a higher trade cost τ_x and dispersion of distortion σ_τ in 1998— at about 43% and 20% higher than the level in 2005. The mean TFP in 2005 is about 45% higher than that in 1998, reflecting technological improvements and factor accumulation over time.

These estimates are then used to run counterfactual experiments, in order to decompose China's growth in between 1998 and 2005. The factors considered include technological progress, input accumulation, and the reduction of trade costs and domestic distortions. In

Table 4: Data, 1998 and 2005

Target Moments	Data (1998)	Data (2005)
Fraction of firms producing ω_e	0.77	0.85
Mean − lowest 5% for $ln(K^{\alpha}L^{1-\alpha})$	2.04	1.82
Fraction of firm exporting	0.25	0.30
Export intensity	0.30	0.41
std of existing firms ln(VA)	1.33	1.20
std of existing firms $ln(VA/K^{\alpha}L^{1-\alpha})$	1.12	0.93
Corr(lnVA, $ln(VA/K^{\alpha}L^{1-\alpha})$)	0.47	0.41
Relative real GDP of US to China	2.50	1.79
Change of China's real GDP		57%

each experiment, the parameters for the year 1998 remain fixed, while each of the following parameters– mean TFP μ_{φ}, trade cost τ_x, or dispersion of distortion σ_{τ}—are allowed to vary to its 2005 level. Table 5 shows that the increase of technology and inputs alone lead to a 44% increase in GDP and a 46% increase in TFP. Reduction in trade costs would independently boost GDP by 8% and TFP by only 3%. In contrast, lowering the dispersion of distortions increases GDP by 66% and TFP by 69%.[22]

A notable point of comparison is with Tombe and Zhu (2019), which, despite adopting an altogether different approach, finds also small gains to trade. In their model that features migration across regions and sectors in China, international trade contributes to only 7% of productivity growth in between 2000 and 2005. In other words, international trade has led to very little allocative benefits of labor across regions and sectors—as compared to direct reforms that lower migration costs or reductions in internal trade costs. Their model does not feature distortions at the firm level that can render trade's allocative benefits even smaller. This leads us to find an even smaller effect of trade on productivity in China over roughly the same period.

Of course, a caveat is that trade may also help reduce domestic distortions. If, say, the WTO requires certain kind of domestic reforms as a pre-condition for entry, then some of the technological improvement and reductions in the level of distortions could be partially induced by opening up policies. We do not consider this here. Also, this quantitative

[22]Note that the contributions to GDP or TFP increase don't add up to 100% because the productivity distribution and fixed costs have also changed from 1998 to 2005. Furthermore, there are interacting effects on mean TFP, trade cost, and distortion dispersions.

exercise of course also ignores other potential channels of gains to trade, such as pro-competition effect of trade, or potentially transfers of technology (Ramondo and Rodríguez-Clare (2013))–though these effects may still be quantitatively small. The point we make here is that in our benchmark framework, the contribution of trade pales in comparison to the contribution of domestic policies and technological progress in accounting for China's growth experience.

Table 5: Decomposition of China's Growth between 1998-2005

	Change of Real GDP	Change of TFP
Benchmark	57%	56%
Counterfactual Change from 1998-2005:		
Technology and inputs alone (Increase mean φ)	44%	46%
Trade alone (Decrease τ_x)	8%	3%
Distortion alone (Decrease σ_τ)	66%	69%

4 Discussion

In this section, we explore a model with endogenous distortion arising from endogenous markup, which has been extensively studied in the standard trade literature. We show that the endogenous markup model runs counter with the data in that exporters in the model face a higher markup and distortion.

We then address the issue of mismeasurement of inputs or outputs. We also use Bils, Klenow, and Ruane (2017)'s method that detect measurement errors and show how it relates to our estimates of marginal product and average product. We show that even taking out the standard measurement errors, there are still large distortions remaining among Chinese firms.

4.1 Endogenous distortions

Here we build a model with endogenous markup as in Edmond, Midrigan, and Xu (2018). The consumer's problem is the same as before.

Final goods producer Final goods producers are competitive and produce with intermediate goods with a Kimball aggregator

$$\int_{\omega \in \Omega} \gamma \left(\frac{q}{Q} \right) d\omega = 1,$$

where $\gamma(\cdot)$ follows Klenow and Willis (2016) specification as

$$\gamma \left(\frac{q}{Q} \right) = 1 + (\sigma - 1) \exp \left(\frac{1}{\varepsilon} \right) \varepsilon^{\frac{\sigma}{\varepsilon} - 1} \left[\Gamma \left(\frac{\sigma}{\varepsilon}, \frac{1}{\varepsilon} \right) - \Gamma \left(\frac{\sigma}{\varepsilon}, \frac{(q/Q)^{\frac{\varepsilon}{\sigma}}}{\varepsilon} \right) \right], \tag{22}$$

$\sigma > 1$, $\varepsilon \geq 0$ and $\Gamma(s, x)$ denotes the upper incomplete Gamma function $\Gamma(s, x) = \int_x^\infty t^{s-1} e^{-t} dt$. The demand function for each intermediate good producer is therefore given by

$$p(\omega) = \gamma' \left(\frac{q(\omega)}{Q} \right) PD, \tag{23}$$

where D is a demand index, $D = \left[\int_{\omega \in \Omega} \gamma' \left(\frac{q(\omega)}{Q} \right) \frac{q(\omega)}{Q} d\omega \right]^{-1}$.

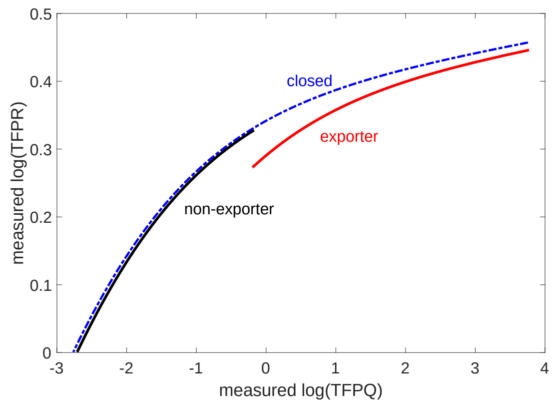
Intermediate good producer The problem of an intermediate good producer is similar as before except it faces a demand function as in equation (23). The firm will choose the price as a markup over the marginal cost,

$$p = \frac{\sigma}{\sigma - (q/Q)^{\frac{\varepsilon}{\sigma}}} \frac{w\tau}{\varphi}.$$

Note that the markup is endogenous and depends on the size of the firm, the higher the quantity a firm sells, the higher the markup it charges. The firm's optimal production and profit increase with φ and decrease with τ. Firms face the same fixed cost and exporting costs as in the Benchmark model, hence there exists a cutoff $\varphi^*(\tau)$, firms produce when $\varphi \geq \varphi^*(\tau)$.

To compare with the benchmark model, we choose ε as 0.08 to match the aggregate marginal product of labor of 1.45 as in Edmond, Midrigan, and Xu (2018), while keeping other parameters the same as in the benchmark. Figure 7 plots the relationship between the measured $\log(TFPR)$ (which again is ARPL, $pq/(\ell_v + f)$ in the model and f includes

Figure 7: Measured TFPR and TFPQ in an Endogenous Markup Model



Notes: TFPQ is measured with $q/(\ell_v + f)$ and TFPR is $pq/(\ell_v + f)$ where ℓ_v is the variable input.

exporting fixed cost if firm exports) and the measured $\log(TFPQ)$ (which is $q/(\ell_v + f)$) in the model). First, higher productivity firms produce more and end up with a higher endogenous markup. The measured TFPR is therefore higher. Hence, we observe an upward sloping line for the closed economy. Second, this upward sloping patterns also show up in the open economy. Moreover, exporters are more productive and face a higher wedge. Non-exporters face a more competitive market after opening up and charge a lower markup, the TFPR is smaller. Around the exporting cutoff, exporters face a lower TFPR due to the fixed cost of exporting. Overall, exporters face higher TFPR.

In summary, if the observed wedges are purely driven by markups and they endogenously change with trade, we should see that: 1) exporters on average have higher markups, hence higher—rather than lower—TFPR; and given TFPQ, they should have the same TFPR; 2) measured log (TFPR) and log (VA) will be almost perfectly correlated. These implications are at odds with the regression results shown in section 3.4, where exporters face lower TFPR. Thus, even in this extended model, similar exogenous distortions are needed to match the observed dispersion and correlation. This is consistent with Song and Wu (2015) and David and Venkateswaran (2017) that the heterogeneity in markup explains very limited MPK dispersion in China. Moreover, Arkolakis et al. (2018) show that the gains from trade in a model with endogenous markup is similar to ACR. In the current extended model with endogenous markup, the excess welfare gain from trade when removing domestic distortions is 3.2%, while it is 5.4% (9.8% − 4.4%) in the benchmark model.

4.2 Detecting measurement error

With the presence of fixed costs in producing and exporting in our model, the measured TFPR does not perfectly relate to the true wedges. In the data, there are other types of mismeasurements in output and input, which may also generate a dispersion in the average revenue products, and thereby affect the measured TFPR— as shown in Bils, Klenow, and Ruane (2017) and Song and Wu (2015). In this section, we address the issues surrounding measurement error following the practices adopted in the literature.

The main approach involves using panel data to estimate the true marginal product dispersion, rather than simply employing cross-sectional data. With this method, we find that the measurement errors are small in China, accounting for only 18% of the variation in the average product.[23] This 18% includes the mismeasurement of production inputs in the presence of fixed cost, which is accounted for in our benchmark.

We exploit three alternative methods to detect measurement error: average annual observations within firms, first differences over years within firms, and covariance between first differences and average products. All three approaches point to the same conclusion: that 1) there is a large dispersion in marginal products in China; 2) measurement error only accounts for a small fraction of the dispersion in the measured marginal products (i.e. average products).

First, if measurement error were idiosyncratic across firms and over time, one can take the time average of annual observations within firms to wash out these errors, drastically reducing the dispersion of average products. The upper panel of Table 6 reports the statistics when we take the average within firms. The average standard deviation is 1.19 for the average product of capital and 0.96 for the average product of labor. The standard deviations of value added and the average product of inputs are 1.19 and 0.94, where the correlation between the two variables is 0.4. These results mimic our benchmark moments.

[23] Bils, Klenow, and Ruane (2017) finds measurement errors can explain about half of variation of average products in Indian, and about 80% of that in the U.S, but little for China.

In particular, the dispersions of average products of inputs are still high. This implies that measurement errors of the iid type cannot explain the observed dispersions in the average products.

Table 6: Detecting Measurement Errors

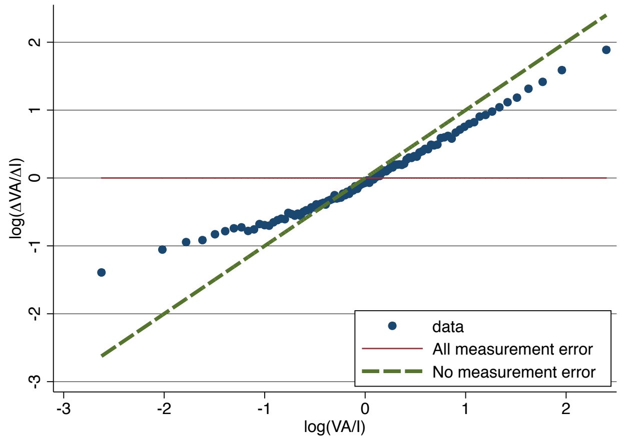
Average annual observation within firm				
std(ln(APK))	std(ln(APL))	$std(lnVA)$	$std(ln(VA/I))$	$corr(lnVA, ln(VA/I))$
1.19	0.96	1.19	0.94	0.4
	First level differences			
		2001	2004	2007
$std(ln(\triangle VA/\triangle K))$		1.82	1.78	1.76
$std(ln(\triangle VA/\triangle L))$		1.68	1.60	1.61
	Regression			
		Ψ	$\Psi(1-\lambda)$	
		0.53^{***}	-0.0997^{***}	
		(34.58)	(−20.65)	

Note: This table reports three ways to detect measurement errors.
The upper panel reports the average annual levels within firms.
The middle panel reports the ratio of first differences as another measure of marginal product, where $\triangle VA$ denotes the first difference of value added.
The lower panel reports regression coefficient as in equation (24).
Robust t-statistics in parentheses.

Second, as pointed out by Bils, Klenow, and Ruane (2017), the dispersion of first differences reflect the true distortion if marginal products are constant over time. Calculating the first differences of value added $\triangle VA$, capital $\triangle K$, and labor $\triangle L$, and then taking the ratio $\triangle VA/\triangle K$ and $\triangle VA/\triangle L$ gives us an alternative measure of marginal products. The 1% tails of both ratios are trimmed, and the results are displayed in the middle panel of Table 6 for the year of 2001, 2004, and 2007. The dispersions are even higher than those in Table A-1 for the measured average product of inputs.

Moreover, the alternative measured marginal products are highly correlated with average products. Figure 8 plots the $\ln(\triangle VA/\triangle I)$ against the benchmark average product of input $\ln(VA/I)$ where I is the composite of inputs, $I = K^{\alpha}L^{1-\alpha}$, where each dot corresponds to one of 100 percentiles of $\ln(VA/I)$. The regression coefficient at the firm level is 0.72, see Table A-3. Note that without measurement errors, the two measures are perfectly correlated. For the case with only measurement error, the two measures have no correlation. Hence, the high correlation between the alternative measure and the average products suggest small measurement errors and a large distortion-induced misallocation.

Figure 8: Measured Marginal Product using First Differences vs TFPR



Lastly, we follow Bils, Klenow, and Ruane (2017) and run the following regression to further quantify the extent to which measured average products reflect marginal products:

$$\triangle \widehat{VA}_i = \Phi \cdot \log(TFPR_i) + \Psi \cdot \triangle \hat{I}_i - \Psi(1-\lambda) \cdot \log(TFPR_i) \cdot \triangle \hat{I}_i + D_s + \xi_i \qquad (24)$$

where $\triangle \widehat{VA}_i$ and $\triangle \hat{I}_i$ are the growth rate of measured value added and inputs respectively, and $\log(TFPR_i)$ is the measured average products. The underlying assumption here is that the measurement errors are additive. The variable of interest in the regression is λ, the variance of distortions relative to that of $TFPR$: $\lambda = \frac{\sigma^2_{\ln \tau}}{\sigma^2_{\ln(TFPR)}}$. The regression coefficient for Ψ is 0.53 and for the interaction of $\log(TFPR_i)$ and $\triangle \hat{I}_i$ is -0.0997. Both are significant, and the robust t-statistics are reported in Table 6. The implied λ is therefore 0.81. Hence, 81% of variation in $TFPR$ or average products is accounted for by distortion τ and 19% is due to measurement errors. The results are robust if we weight the observations with their share of aggregate value added or if we control for higher orders of $\ln(TFPR)$ to allow for stationary shocks to firms productivity and distortions. See the Appendix E for details.[24]

In summary, the three alternative ways of sifting out measurement errors using panel data all point to the result that the dispersion in the average product of inputs are mainly driven by distortions rather than measurement error typically conceived.

[24] Bils, Klenow, and Ruane (2017) also considers the following extension to allow for stationary shocks to

5 Conclusion

This paper evaluates the impact of trade liberalization when the economy is subject to firm-level distortions. Given its prevalence and importance in developing countries, it is reasonable to ask how trade might affect welfare when these distortions are taken into account. This paper shows theoretically and quantitatively that opening an economy may in fact reduce allocative efficiency and exacerbate the misallocation of resources, by helping firms that are more subsidized (rather than those who are more productive) to expand. The findings in these papers, nevertheless, in no way disclaims the potential wide variety of sources and the magnitude of gains to trade beyond what is conventionally modelled and taken up in the current framework. But it does highlight that these losses could be sizeable and comparable to major sources of welfare gains. We use Chinese manufacturing data in a period of the economy's rapid integration to demonstrate quantitatively that standard calculations for welfare may grossly overestimate the gains.

The paper serves as a first attempt to understand the interactions between trade and idiosyncratic firm level distortions on a theoretical level. The assumption of exogenous distortions is both reasonable empirically but also helpful in laying bare the key theoretical mechanisms on their interaction with trade. Of course, the next step would be to examine more thoroughly how trade might partially change the distribution and nature of these distortions. Extensions of the work are inherently numerous and promisingly fruitful. One can examine how distortions interact with other channels of gains to trade, such as innovation. One can also examine a dynamic model and the sequence of trade and domestic reforms. Our work joins the growing body of work and interest on why developing countries' experience with trade liberalization might have been so curiously diverse and uneven. Our work hopefully lends itself as one explanation to such a question.

firms productivity and distortions:

$$\begin{aligned}\triangle\widehat{VA}_i &= \Phi\cdot\log(TFPR_i) + \Psi\cdot\triangle\hat{I}_i - \Psi(1-\lambda)\cdot\log(TFPR_i)\\ &+ \Gamma\cdot[\log(TFPR_i)]^2 + \Lambda(1-\lambda)\cdot[\log(TFPR_i)]^2\triangle\hat{I}_i\\ &+ \Upsilon\cdot[\log(TFPR_i)]^3 + \Lambda(1-\lambda)\cdot[\log(TFPR_i)]^3\triangle\hat{I}_i.\end{aligned}$$

References

Alvarez, Fernando E, Francisco J Buera, and Robert E Lucas Jr. 2013. "Idea flows, economic growth, and trade." Tech. rep., National Bureau of Economic Research.

Anderson, James E and Eric Van Wincoop. 2004. "Trade costs." *Journal of Economic literature* 42 (3):691–751.

Arkolakis, Costas, Arnaud Costinot, Dave Donaldson, and Andrés Rodríguez-Clare. 2018. "The elusive pro-competitive effects of trade." *The Review of Economic Studies* 86 (1):46–80.

Arkolakis, Costas, Arnaud Costinot, and Andrés Rodríguez-Clare. 2012. "New trade models, same old gains?" *American Economic Review* 102 (1):94–130.

Asturias, Jose, Sewon Hur, Timothy J Kehoe, and Kim J Ruhl. 2016. "The interaction and sequencing of policy reforms." *Journal of Economic Dynamics and Control* 72:45–66.

Atkeson, Andrew and Ariel Tomas Burstein. 2010. "Innovation, firm dynamics, and international trade." *Journal of political economy* 118 (3):433–484.

Bai, Chong-En, Chang-Tai Hsieh, and Zheng Michael Song. 2019. "Special deals with chinese characteristics." Tech. rep., National Bureau of Economic Research.

Bai, Yan, Keyu Jin, and Dan Lu. 2020. "How different domestic frictions affect trade." Tech. rep., working paper.

Bai, Yan, Dan Lu, and Xu Tian. 2018. "Do Financial Frictions Explain Chinese Firms? Saving and Misallocation?" Tech. rep., National Bureau of Economic Research.

Baily, Martin N, Charles Hulten, and David Campbell. 1992. "The distribution of productivity in manufacturing plants." *Brookings Papers: Microeconomics* 187:249.

Baqaee, David Rezza and Emmanuel Farhi. 2019. "Networks, Barriers, and Trade." Tech. rep., National Bureau of Economic Research.

———. 2020. "Productivity and Misallocation in General Equilibrium." Tech. rep., Quarterly Journal of Economics.

Bartelsman, Eric, John Haltiwanger, and Stefano Scarpetta. 2009. "Measuring and analyzing cross-country differences in firm dynamics." In *Producer dynamics: New evidence from micro data*. University of Chicago Press, 15–76.

Bernard, Andrew B, Jonathan Eaton, J Bradford Jensen, and Samuel Kortum. 2003. "Plants and productivity in international trade." *American economic review* 93 (4):1268–1290.

Berthou, Antoine, John Jong&Hyun Chung, Kalina Manova, and Charlotte Sandoz Dit Bragard. 2018. "Productivity,(mis) allocation and trade." Tech. rep., Mimeo.

Bhagwati, Jagdish and Vangal K Ramaswami. 1963. "Domestic distortions, tariffs and the theory of optimum subsidy." *Journal of Political economy* 71 (1):44–50.

Bils, Mark, Peter J. Klenow, and Cian Ruane. 2017. "Misallocation or Mismeasurement?" Working paper, University of Rochester.

Brandt, Loren, Johannes Van Biesebroeck, and Yifan Zhang. 2012. "Creative accounting or creative destruction? Firm-level productivity growth in Chinese manufacturing." *Journal of development economics* 97 (2):339–351.

Buera, Francisco J and Ezra Oberfield. 2016. "The global diffusion of ideas." Tech. rep., National Bureau of Economic Research.

Burstein, Ariel and Javier Cravino. 2015. "Measured aggregate gains from international trade." *American Economic Journal: Macroeconomics* 7 (2):181–218.

Chang, Roberto, Linda Kaltani, and Norman V Loayza. 2009. "Openness can be good for growth: The role of policy complementarities." *Journal of development economics* 90 (1):33–49.

Chen, Ting and James Kai-sing Kung. 2018. "Busting the "Princelings": The Campaign Against Corruption in China's Primary Land Market." *The Quarterly Journal of Economics* 134 (1):185–226.

Costa-Scottini, Lucas. 2018. "Firm-Level Distortions, Trade, and International Productivity Differences." Working paper, Brown University.

David, Joel M and Venky Venkateswaran. 2017. "The sources of capital misallocation." Tech. rep., National Bureau of Economic Research.

Edmond, Chris, Virgiliu Midrigan, and Daniel Yi Xu. 2018. "How costly are markups?" Tech. rep., National Bureau of Economic Research.

Guo, D, K Jiang, BY Kim, and C Xu. 2013. "The political economy of private firms in China." *Journal of Finance* 61 (6):2597–2635.

Harrison, Ann and Andrés Rodríguez-Clare. 2010. "Trade, foreign investment, and industrial policy for developing countries." In *Handbook of development economics*, vol. 5. Elsevier, 4039–4214.

Ho, Giang T. 2010. "Trade Liberalization with Idiosyncratic Distortions: Theory and Evidence from India." Working paper.

Hsieh, Chang-Tai and Peter J Klenow. 2009. "Misallocation and manufacturing TFP in China and India." *The Quarterly journal of economics* 124 (4):1403–1448.

Hsieh, Chang-Tai, Nicholas Li, Ralph Ossa, and Mu-Jeung Yang. 2016. "Accounting for the new gains from trade liberalization." Tech. rep., National Bureau of Economic Research.

Khandelwal, Amit K, Peter K Schott, and Shang-Jin Wei. 2013. "Trade liberalization and embedded institutional reform: Evidence from Chinese exporters." *American Economic Review* 103 (6):2169–95.

Klenow, Peter J and Jonathan L Willis. 2016. "Real rigidities and nominal price changes." *Economica* 83 (331):443–472.

Krugman, Paul R, Maurice Obstfeld, and Marc J Melitz. 2015. "International economics: theory and policy."

Melitz, Marc J. 2003. "The impact of trade on intra-industry reallocations and aggregate industry productivity." *econometrica* 71 (6):1695–1725.

Melitz, Marc J and Stephen J Redding. 2015. "New trade models, new welfare implications." *American Economic Review* 105 (3):1105–46.

Perla, Jesse, Christopher Tonetti, and Michael E Waugh. 2015. "Equilibrium technology diffusion, trade, and growth." Tech. rep., National Bureau of Economic Research.

Ramondo, Natalia and Andrés Rodríguez-Clare. 2013. "Trade, multinational production, and the gains from openness." *Journal of Political Economy* 121 (2):273–322.

Restuccia, Diego and Richard Rogerson. 2008. "Policy distortions and aggregate productivity with heterogeneous establishments." *Review of Economic dynamics* 11 (4):707–720.

Rodríguez-Clare, Andrés. 2018. *Globalization, And the Gains from Trade in Rich and Poor Countries*. Princeton University Press.

Sampson, Thomas. 2015. "Dynamic selection: an idea flows theory of entry, trade, and growth." *The Quarterly Journal of Economics* 131 (1):315–380.

Simonovska and Waugh. 2014. "Trade Models, Trade Elasticities, and the Gains from Trade." .

Song, Zheng and Guiying Laura Wu. 2015. "Identifying capital misallocation." Tech. rep., Working paper.

Tombe, Trevor and Xiaodong Zhu. 2019. "Trade, migration, and productivity: A quantitative analysis of china." *American Economic Review* 109 (5):1843–72.

Waugh, Michael E. 2010. "International trade and income differences." *American Economic Review* 100 (5):2093–2124.

Wu, Guiying Laura. 2018. "Capital misallocation in China: Financial frictions or policy distortions?" *Journal of Development Economics* 130:203–223.

Online Appendix to "Misallocation Under Trade Liberalization"

by Yan Bai, Keyu Jin, and Dan Lu

A Model Derivation

Closed Economy Equilibrium. In a closed economy, the free entry condition requires that the present value of producing equals the entry cost. The probability of entry ω_e is given by

$$\omega_e = \int \int_{\varphi^*(\tau)}^{\infty} g(\varphi, \tau)\, d\varphi d\tau, \tag{A.1}$$

and the distribution of operating firms $\mu(\varphi, \tau)$

$$\mu(\varphi, \tau) = \frac{g(\varphi, \tau)}{\int \int_{\varphi^*(\tau)}^{\infty} g(\varphi, \tau)\, d\varphi d\tau} = \frac{g(\varphi, \tau)}{\omega_e}$$

if $\varphi \geq \varphi^*(\tau)$; and 0 otherwise. Let the per-period expected profit conditional on producing be

$$E[\pi(\varphi, \tau)] = \int \int_{\varphi^*(\tau)}^{\infty} \pi(\varphi, \tau) \mu(\varphi, \tau) d\varphi d\tau.$$

The free entry condition is given by

$$\omega_e E[\pi(\varphi, \tau)] = w f_e. \tag{A.2}$$

The optimal profit function (3) combined with the above expression in the free entry condition yields an equation for P,

$$\frac{PQ}{\sigma}\left(P\frac{\sigma-1}{\sigma}\right)^{\sigma-1} w^{1-\sigma} \int \int_{\varphi^*(\tau)} \left[\varphi^{\sigma-1}\tau^{-\sigma}\right] g(\varphi, \tau) d\varphi d\tau - wf \int \int_{\varphi^*(\tau)} g(\varphi, \tau) d\varphi d\tau = w f_e. \tag{A.3}$$

Let M_e be the measure of new entrant. The labor market clearing condition requires

$$L = ME\left[\frac{q}{\varphi} + f\right] + M_e f_e,$$

where the average labor demanded by firms $E\left[\frac{q}{\varphi}+f\right]$ is given by

$$E\left[\frac{q}{\varphi}+f\right]=\int\int_{\varphi^{*}(\tau)}^{\infty}\left[\frac{q}{\varphi}+f\right]\mu\left(\varphi,\tau\right)d\varphi d\tau.$$

In a stationary equilibrium, the number of entrants equals number of exits, such that $\omega_e M_e = M$ and

$$M=\frac{L}{\sigma\left(\frac{f_e}{\omega_e}+f\right)}. \tag{A.4}$$

Open Economy Equilibrium. Optimal prices and cutoff functions are straightforward analogues of the closed economy case. The free entry condition for Home and Foreign implies

$$\begin{aligned}&\frac{PQ}{\sigma}\left(P\frac{\sigma-1}{\sigma}\right)^{\sigma-1}w^{1-\sigma}\int\int_{\varphi^{*}(\tau)}^{\infty}\left[\varphi^{\sigma-1}\tau^{-\sigma}\right]g\left(\varphi,\tau\right)d\varphi d\tau-wf\int\int_{\varphi^{*}(\tau)}^{\infty}g\left(\varphi,\tau\right)d\varphi d\tau\\&+\left[\frac{P_fQ_f}{\sigma}\left(P_f\frac{\sigma-1}{\sigma}\right)^{\sigma-1}(\tau_x w)^{1-\sigma}\int\int_{\varphi_x^{*}(\tau)}^{\infty}\left[\varphi^{\sigma-1}\tau^{-\sigma}\right]g\left(\varphi,\tau\right)d\varphi d\tau-wf_x\int\int_{\varphi_x^{*}(\tau)}^{\infty}g\left(\varphi,\tau\right)d\varphi d\tau\right]\\&=wf_e. \qquad \text{(A.5)}\end{aligned}$$

$$\begin{aligned}&\frac{P_fQ_f}{\sigma}\left(P_f\frac{\sigma-1}{\sigma}\right)^{\sigma-1}w_f^{1-\sigma}\int\int_{\varphi_f^{*}(\tau)}^{\infty}\left[\varphi^{\sigma-1}\tau^{-\sigma}\right]g_f\left(\varphi,\tau\right)d\varphi d\tau-w_ff\int\int_{\varphi_f^{*}(\tau)}^{\infty}g_f\left(\varphi,\tau\right)d\varphi d\tau\\&+\left[\frac{PQ}{\sigma}\left(P\frac{\sigma-1}{\sigma}\right)^{\sigma-1}(\tau_x w_f)^{1-\sigma}\int\int_{\varphi_{xf}^{*}(\tau)}^{\infty}\varphi^{\sigma-1}\tau^{-\sigma}g_f\left(\varphi,\tau\right)d\varphi d\tau-w_ff_x\int\int_{\varphi_{xf}^{*}(\tau)}^{\infty}g_f\left(\varphi,\tau\right)d\varphi d\tau\right]\\&=w_ff_e. \qquad \text{(A.6)}\end{aligned}$$

It follows that the aggregate prices of Home and Foreign are

$$P^{1-\sigma}=\left(\frac{\sigma}{\sigma-1}\right)^{1-\sigma}\left[Mw^{1-\sigma}\frac{\int\int_{\varphi^{*}(\tau)}^{\infty}(\frac{\varphi}{\tau})^{\sigma-1}g\left(\varphi,\tau\right)d\varphi d\tau}{\int\int_{\varphi^{*}(\tau)}^{\infty}g\left(\varphi,\tau\right)d\varphi d\tau}+M_f(\tau_xw_f)^{1-\sigma}\frac{\int\int_{\varphi_{xf}^{*}(\tau)}^{\infty}(\frac{\varphi}{\tau})^{\sigma-1}g_f\left(\varphi,\tau\right)d\varphi d\tau}{\int\int_{\varphi_f^{*}(\tau)}^{\infty}g_f\left(\varphi,\tau\right)d\varphi d\tau}\right]$$

$$P_f^{1-\sigma}=\left(\frac{\sigma}{\sigma-1}\right)^{1-\sigma}\left[M_fw_f^{1-\sigma}\frac{\int\int_{\varphi^{*}(\tau)_f}^{\infty}(\frac{\varphi}{\tau})^{\sigma-1}g_f\left(\varphi,\tau\right)d\varphi d\tau}{\int\int_{\varphi_f^{*}(\tau)}^{\infty}g_f\left(\varphi,\tau\right)d\varphi d\tau}+M(w\tau_x)^{1-\sigma}\frac{\int\int_{\varphi^{*}(\tau)_x}^{\infty}(\frac{\varphi}{\tau})^{\sigma-1}g\left(\varphi,\tau\right)d\varphi d\tau}{\int\int_{\varphi^{*}(\tau)}^{\infty}g\left(\varphi,\tau\right)d\varphi d\tau}\right].$$

Next we derive the measure of operating firms at Home and Foreign M and M_f. Let M_e be the measure of new entrant, ω_e be the entry probability given by (A.1), and ω_x be the export probability conditional on entry

$$\omega_x = \int\int_{\varphi_x^*(\tau)}^{\infty} \mu(\varphi,\tau)\, d\varphi d\tau = \frac{\int\int_{\varphi_x^*(\tau)}^{\infty} g(\varphi,\tau)\, d\varphi d\tau}{\int\int_{\varphi^*(\tau)}^{\infty} g(\varphi,t)\, d\varphi d\tau}.$$

The expected per-period profit includes the profit from both domestic production and international production,

$$\begin{aligned}&\int\int_{\varphi^*(\tau)} \pi(\varphi,\tau)\mu(\varphi,\tau)d\varphi d\tau + \int\int_{\varphi_x^*(\tau)} \pi_x(\varphi,\tau)\mu(\varphi,\tau)d\varphi d\tau \\ &\equiv E\pi + \omega_x E\pi_x\end{aligned}$$

where $\mu_x(\varphi,\tau) = \mu(\varphi,\tau)/\omega_x$. Using the average profits, free entry, stationary equilibrium condition, we have

$$M = \frac{L}{\sigma\left(\frac{f_e}{\int\int_{\varphi^*(\tau)}^{\infty} g(\varphi,\tau)d\varphi d\tau} + f + \frac{\int\int_{\varphi_x^*(\tau)}^{\infty} g(\varphi,\tau)d\varphi d\tau}{\int\int_{\varphi^*(\tau)}^{\infty} g(\varphi,\tau)d\varphi d\tau} f_x\right)} \tag{A.7}$$

and

$$M_f = \frac{L_f}{\sigma\left(\frac{f_e}{\int\int_{\varphi_f^*(\tau)}^{\infty} g_f(\varphi,\tau)d\varphi d\tau} + f + \frac{\int\int_{\varphi_{xf}^*(\tau)}^{\infty} g_f(\varphi,\tau)d\varphi d\tau}{\int\int_{\varphi_f^*(\tau)}^{\infty} g_f(\varphi,\tau)d\varphi d\tau} f_x\right)}. \tag{A.8}$$

Finally the assumption of a balanced trade results in

$$P_f^{\sigma} Q_f M \int\int_{\varphi_x^*(\tau)}^{\infty} \left(\frac{w\tau_x\tau}{\varphi}\right)^{1-\sigma} \mu(\varphi,\tau)\, d\varphi d\tau = P^{\sigma} Q M_f \int\int_{\varphi_{xf}^*(\tau)}^{\infty} \left(\frac{w_f\tau_x\tau}{\varphi}\right)^{1-\sigma} \mu_f(\varphi,\tau)\, d\varphi d\tau. \tag{A.9}$$

B Proofs

B.1 Proof for Proposition 2

Proof. (1) In a closed economy, from the labor market clearing condition,

$$M_e \int \int_{\varphi^*(\tau)} (\frac{\varphi}{\tau})^{\sigma-1} \frac{1}{\tau} P^{\sigma-1} PQg(\varphi, \tau) d\varphi d\tau = L,$$

we get

$$d \ln W = d \ln Q = -d \ln P - d \ln M_e - (\sigma - 1) d \ln P - d \ln \int \int_{\varphi^*(\tau)} (\frac{\varphi}{\tau})^{\sigma-1} \frac{1}{\tau} g(\varphi, \tau) d\varphi d\tau. \tag{A.10}$$

From the price index,

$$P^{1-\sigma} = M_e \int \int_{\varphi^*(\tau)} (\frac{\varphi}{\tau})^{\sigma-1} g(\varphi, \tau) d\varphi d\tau,$$

$$\begin{aligned}(1-\sigma) d \ln P &= d \ln M_e + \frac{\int (\sigma-1)(\frac{\varphi_i}{\tau})^{\sigma-1} g(\varphi_i, \tau) d\tau}{\int \int_{\varphi^*(\tau)} (\frac{\varphi}{\tau})^{\sigma-1} g(\varphi, \tau) d\varphi d\tau} d \ln \varphi_i - \gamma_\lambda(\hat{\varphi}^*) d \ln \hat{\varphi}^* \\ &= d \ln M_e + (\sigma - 1) \frac{p_i q_i}{PQ} d \ln \varphi_i - \gamma_\lambda(\hat{\varphi}^*) d \ln \hat{\varphi}^*\end{aligned}$$

where

$$p_i q_i = M_e \int (\frac{\varphi_i}{\tau})^{\sigma-1} P^\sigma Q g(\varphi_i, \tau) d\tau,$$

Substitute $(1-\sigma) d \ln P$ into equation (A.10):

$$d \ln W = -d \ln P + (\sigma - 1)[\frac{p_i q_i}{PQ} - \frac{\ell_i}{L}] d \ln \varphi_i + (\gamma_s - \gamma_\lambda) d \ln \hat{\varphi}^*.$$

where

$$\ell_i = M_e \int (\frac{\varphi_i}{\tau})^{\sigma-1} \frac{1}{\tau} P^\sigma Q g(\varphi_i, \tau) d\tau.$$

(2) To derive the effect of trade cost shock, we use the following equilibrium equations,

(a) Free entry condition:

$$w^{1-\sigma}\left[P^{\sigma}Q\int\int_{\varphi^*(\tau)}\varphi^{\sigma-1}\tau^{-\sigma}g(\varphi,\tau)d\varphi d\tau+P_f^{\sigma}Q_f\tau_x^{1-\sigma}\int\int_{\varphi_x^*(\tau)}\varphi^{\sigma-1}\tau^{-\sigma}g(\varphi,\tau)d\varphi d\tau\right]$$
$$=\sigma^{\sigma}(\sigma-1)^{1-\sigma}\left(wf_e+wfH+wf_xH_x\right)$$

where $H=\int\int_{\varphi^*(\tau)}g(\varphi,\tau)d\varphi d\tau$ and $H_x=\int\int_{\varphi_x^*(\tau)}g(\varphi,\tau)d\varphi d\tau$.

(b) The labor market clearing condition:

$$M_e=\frac{L}{\sigma\left(f_e+Hf+H_xf_x\right)}.$$

(c) Price index:

$$P^{1-\sigma}=con_p\times\left[M\frac{\int\int_{\varphi^*(\tau)}^{\infty}(\frac{\varphi}{\tau})^{\sigma-1}g\left(\varphi,\tau\right)d\varphi d\tau}{w^{\sigma-1}\int\int_{\varphi^*(\tau)}^{\infty}g\left(\varphi,\tau\right)d\varphi d\tau}+M_f\frac{\int\int_{\varphi_{xf}^*(\tau)}^{\infty}(\frac{\varphi}{\tau})^{\sigma-1}g_f\left(\varphi,\tau\right)d\varphi d\tau}{\left(\tau_xw_f\right)^{\sigma-1}\int\int_{\varphi_f^*(\tau)}^{\infty}g_f\left(\varphi,\tau\right)d\varphi d\tau}\right]$$

combines with trade balance:

$$P_f^{\sigma}Q_fM\int\int_{\varphi_x^*(\tau)}^{\infty}\left(\frac{w\tau_x\tau}{\varphi}\right)^{1-\sigma}\mu\left(\varphi,\tau\right)d\varphi d\tau=P^{\sigma}QM_f\int\int_{\varphi_{xf}^*(\tau)}^{\infty}\left(\frac{w_f\tau_x\tau}{\varphi}\right)^{1-\sigma}\mu_f\left(\varphi,\tau\right)d\varphi d\tau.$$

we get

$$P^{1-\sigma}=con_pM_ew^{1-\sigma}\left[\int\int_{\varphi^*(\tau)}(\frac{\varphi}{\tau})^{\sigma-1}g(\varphi,\tau)d\varphi d\tau+\frac{P_f^{\sigma}Q_f}{P^{\sigma}Q}\tau_x^{1-\sigma}\int\int_{\varphi_x^*(\tau)}(\frac{\varphi}{\tau})^{\sigma-1}g(\varphi,\tau)d\varphi d\tau\right],$$

where con_p is a constant that depends on the model parameters.

(d) Cutoff of producing:

$$\varphi^*(\tau)=con_v\times P^{-1}\left(PQ\right)^{\frac{1}{1-\sigma}}\tau^{\frac{\sigma}{\sigma-1}}$$

where con_v is a constant that depends on the model parameters. $\hat{\varphi}=\varphi\tau^{\frac{\sigma}{1-\sigma}}$ and $\hat{\varphi}^*=con_v\times P^{-1}\left(PQ\right)^{\frac{1}{1-\sigma}}$.

(e) Definition of domestic share λ:

$$\frac{1-\lambda}{\lambda} = \frac{\frac{P_f^\sigma Q_f}{P^\sigma Q}\tau_x^{1-\sigma}\left[\int\int_{\varphi_x^*(\tau)}\varphi^{\sigma-1}\tau^{1-\sigma}g(\varphi,\tau)d\varphi d\tau\right]}{\left[\int\int_{\varphi^*(\tau)}\varphi^{\sigma-1}\tau^{1-\sigma}g(\varphi,\tau)d\varphi d\tau\right]}$$

(f) Definition of the share of variable labor used in producing domestic goods S:

$$\frac{1-S}{S} = \frac{\frac{P_f^\sigma Q_f}{P^\sigma Q}\tau_x^{1-\sigma}\int\int_{\varphi_x^*(\tau)}\varphi^{\sigma-1}\tau^{-\sigma}g(\varphi,\tau)d\varphi d\tau}{\int\int_{\varphi^*(\tau)}\varphi^{\sigma-1}\tau^{-\sigma}g(\varphi,\tau)d\varphi d\tau}$$

Define O to be the cumulative sales share in the domestic market,

$$O(\hat{\varphi}) = \frac{\int\int_0^{\hat{\varphi}\tau^{\sigma/(\sigma-1)}}\varphi^{\sigma-1}\tau^{1-\sigma}g(\varphi,\tau)d\varphi d\tau}{\int\int_0^\infty\varphi^{\sigma-1}\tau^{1-\sigma}g(\varphi,\tau)d\varphi d\tau}.$$

Similarly define I the cumulative variable input share in the domestic market,

$$I(\hat{\varphi}) = \frac{\int\int_0^{\hat{\varphi}\tau^{\sigma/(\sigma-1)}}\varphi^{\sigma-1}\tau^{-\sigma}g(\varphi,\tau)d\varphi d\tau}{\int\int_0^\infty\varphi^{\sigma-1}\tau^{-\sigma}g(\varphi,\tau)d\varphi d\tau}.$$

The elasticities below are also known as the hazard functions. The elasticity related to variable inputs, or the hazard function for the distribution of log labor size, which is also proportional to the distribution of firm after-tax sales within the domestic market, is given by $\gamma_s(\hat{\varphi}) = -\frac{d\ln(1-I(\hat{\varphi}))}{d\ln\hat{\varphi}}$, and $\varphi^*(\tau) = \hat{\varphi}^*\tau^{\sigma/(\sigma-1)}$,

$$\gamma_s(\hat{\varphi}^*) = \frac{\int\varphi^*(\tau)^{\sigma-1}\tau^{-\sigma}g(\varphi^*(\tau),\tau)\varphi^*(\tau)d\tau}{\int\int_{\varphi^*(\tau)}\varphi^{\sigma-1}\tau^{-\sigma}g(\varphi,\tau)d\varphi d\tau}.$$

The elasticity related to sales, or the hazard function of the distribution of log sales within the domestic market, is given by $\gamma_\lambda = -\frac{d\ln(1-O(\hat{\varphi}))}{d\hat{\varphi}}$, and

$$\gamma_\lambda(\hat{\varphi}^*) = \frac{\int\left(\frac{\varphi^*(\tau)}{\tau}\right)^{\sigma-1}g(\varphi^*(\tau),\tau)\varphi^*(\tau)d\tau}{\int\int_{\varphi^*(\tau)}(\frac{\varphi}{\tau})^{\sigma-1}g(\varphi,\tau)d\varphi d\tau}.$$

Let γ with a superscript x refer to the corresponding elasticity within foreign market:

$$\gamma_\lambda^x = \frac{\int \left(\varphi_x^*(\tau)\right)^{\sigma-1}\tau^{1-\sigma} g(\varphi_x^*(\tau),\tau)\varphi_x^*(\tau) d\tau}{\int \int_{\varphi_x^*(\tau)} \varphi^{\sigma-1}\tau^{1-\sigma} g(\varphi,\tau) d\varphi d\tau}.$$

Differentiating the above system of equations and combining items

$$\begin{aligned} d\ln(PQ) &= (1-\sigma)d\ln P - d\ln M_e + \left(S\gamma_s + S_x\gamma_s^x\right) d\ln\hat{\varphi}^* - S_x\left(1-\sigma-\gamma_s^x\right) d\ln\tau_x \\ &\quad + S_x(1+\frac{\gamma_s^x}{\sigma-1})d\ln\frac{P^\sigma Q}{P_f^\sigma Q_f} \end{aligned} \tag{A.11}$$

$$(1-\sigma-\gamma_\lambda)d\ln P = d\ln M_e - d\ln\lambda + \frac{\gamma_\lambda}{\sigma-1}d\ln(PQ) \tag{A.12}$$

$$d\ln\hat{\varphi}^* = -d\ln P - \frac{1}{\sigma-1}d\ln(PQ) \tag{A.13}$$

$$\begin{aligned} -d\ln\lambda &= (1-\lambda)(1-\sigma-\gamma_\lambda^x)d\ln\tau_x + (1-\lambda)(\gamma_\lambda - \gamma_\lambda^x)d\ln\hat{\varphi}^* \\ &\quad - (1-\lambda)(1+\frac{\gamma_\lambda^x}{\sigma-1})d\ln\frac{P^\sigma Q}{P_f^\sigma Q_f} \end{aligned} \tag{A.14}$$

$$\begin{aligned} -d\ln S &= (1-S)(1-\sigma-\gamma_s^x)d\ln\tau_x + (1-S)(\gamma_s - \gamma_s^x)d\ln\hat{\varphi}^* \\ &\quad - (1-S)(1+\frac{\gamma_s^x}{\sigma-1})d\ln\frac{P^\sigma Q}{P_f^\sigma Q_f} \end{aligned} \tag{A.15}$$

Because

$$(1-\sigma)d\ln P = d\ln M_e - d\ln\lambda - \gamma_\lambda d\ln\hat{\varphi}^* \tag{A.16}$$

$$d\ln(PQ) = (1-\sigma)d\ln P - d\ln M_e + d\ln S + \gamma_s d\ln\hat{\varphi}^* \tag{A.17}$$

We get

$$d\ln Q = -d\ln P + (-d\ln\lambda + d\ln S) + (\gamma_s - \gamma_\lambda)d\ln\hat{\varphi}^*.$$

□

B.2 Proof for Proposition 3

Proof. Solving the system of equations (A.11-A.15), we get Proposition 3

$$
\begin{aligned}
d\ln Q = &-\frac{1}{\gamma_s+\sigma-1}d\ln\lambda \qquad (ACR) \\
&+\frac{1}{\gamma_s+\sigma-1}d\ln M_e \qquad (Redding) \\
&+\frac{1}{\gamma_s+\sigma-1}\frac{\sigma}{\sigma-1}(\gamma_s-\gamma_\lambda)d\ln M_e \\
&+\frac{1}{\gamma_s+\sigma-1}\left[-\left(\sigma-1+\frac{\sigma\gamma_s}{\sigma-1}\right)d\ln\lambda+\left(\sigma-1+\frac{\sigma\gamma_\lambda}{\sigma-1}\right)d\ln S\right]
\end{aligned}
\tag{A.18}
$$

1. Without domestic distortions, $S=\lambda$ and $\gamma_s=\gamma_\lambda$. If productivity follows a Pareto distribution with parameter θ, $\gamma_\lambda=\theta-\sigma+1$ and $dlnM_e=0$. Hence, $dlnW=\frac{1}{\theta}[-dln\lambda]$ as in ACR.

2. Under a general distribution function and without domestic distortions, $S=\lambda$, $\gamma_s=\gamma_\lambda$, but they are not constant, and $dlnM_e\neq 0$. The reallocation term is zero, hence, $dlnW=\frac{1}{\gamma_\lambda+\sigma-1}[-dln\lambda+dlnM_e]$. Micro structure matters for $\gamma_\lambda(\varphi^*)$ and welfare as in MR.

3. With homogenous productivity and if domestic distortion $1/\tau$ follows a Pareto distribution with parameter θ, $\gamma_\lambda=\frac{\sigma-1}{\sigma}(\theta-\sigma+1)$ and $\gamma_s=\frac{\sigma-1}{\sigma}(\theta-\sigma)$, hence $dlnW=\frac{\sigma}{\sigma-1}[dlnS-dln\lambda]$. Changes in the share of expenditure on domestic goods and the share of variable labor used in producing domestic goods are sufficient statistics for welfare change.

□

B.3 Proof for Corollary 1

Proof. Under the special case, $\gamma_\lambda=\frac{\sigma-1}{\sigma}(\theta-\sigma+1)$ and $\gamma_s=\frac{\sigma-1}{\sigma}(\theta-\sigma)$, and the change of welfare becomes $dlnW=\frac{\sigma}{\sigma-1}[dlnS-dln\lambda]$.

1. Welfare change from close to open:

Because domestic shares are

$$\lambda=\left[\frac{P_f^\sigma Q_f}{P^\sigma Q}\tau_x^{1-\sigma}\left(\frac{\tau_x^{\sigma-1}f_x}{f}\frac{P^\sigma Q}{P_f^\sigma Q_f}\right)^{\frac{\sigma-\gamma-1}{\sigma}}+1\right]^{-1}$$

$$S = \left[\frac{P_f^\sigma Q_f}{P^\sigma Q} \tau_x^{1-\sigma} (\frac{\tau_x^{\sigma-1} f_x}{f} \frac{P^\sigma Q}{P_f^\sigma Q_f})^{\frac{\sigma-\gamma}{\sigma}} + 1 \right]^{-1},$$

as long as there is selection to export, i.e., $\frac{\tau_x^{\sigma-1} f_x}{f} \frac{P^\sigma Q}{P_f^\sigma Q_f} > 1$, $\lambda > S$. Intuitively, with reallocation purely driven by distortions, in the open economy, the input share used to produce for exports exceeds the export share. $d \ln S$ is more negative than $d \ln \lambda$ from closed to open. Hence the open economy has an unambiguously lower welfare than in the closed economy.

2. The reallocation term is always negative:

In the welfare expression of Prop 2, the second and third terms cancel out, and the welfare change becomes $\frac{\sigma}{(\sigma-1)\theta}[-d \ln \lambda - \left(\sigma - 1 + \frac{\sigma\gamma_s}{\sigma-1}\right) d \ln \lambda + \left(\sigma - 1 + \frac{\sigma\gamma_\lambda}{\sigma-1}\right) d \ln S] = \frac{\sigma}{(\sigma-1)\theta}[-d \ln \lambda + (1-\theta) d \ln \lambda + \theta d \ln S]$, where the reallocation term is $(1-\theta) d \ln \lambda + \theta d \ln S$.

$$d \ln \lambda = -(1-\lambda) \left(1 - \sigma - \frac{\sigma-1}{\sigma}(\theta - \sigma + 1)\right) d \ln \tau_x + \frac{(1-\lambda)(1-\sigma-\gamma_\lambda^x)}{\sigma-1} d \ln(P_f^\sigma Q_f / P^\sigma Q)$$

$$= (1-\lambda) \frac{\theta+1}{\sigma} d \ln \frac{\tau_x^{\sigma-1} P^\sigma Q}{P_f^\sigma Q_f}$$

$$d \ln S = -(1-S) \left(1 - \sigma - \frac{\sigma-1}{\sigma}(\theta - \sigma)\right) d \ln \tau_x + \frac{(1-S)(1-\sigma-\gamma_s^x)}{\sigma-1} d \ln(P_f^\sigma Q_f / P^\sigma Q)$$

$$= (1-S) \frac{\theta}{\sigma} d \ln \frac{\tau_x^{\sigma-1} P^\sigma Q}{P_f^\sigma Q_f}$$

Substitute for $d \ln \lambda$ and $d \ln S$, the reallocation term is

$$(1-\theta) d \ln \lambda + \theta d \ln S = \frac{\theta^2(1-S) - (\theta^2-1)(1-\lambda)}{\sigma} d \ln \frac{\tau_x^{\sigma-1} P^\sigma Q}{P_f^\sigma Q_f}$$

Substitute for λ and S, it can be shown that $\theta^2(1-S) - (\theta^2-1)(1-\lambda) > 0$, hence as long as the trade cost reduction induces larger fraction of exporters, the reallocation term is always negative. Q.E.D. □

B.4 Proof for Corollary 2

Proof. Recall the producing cutoff is given by $\varphi^*(\tau) = \hat{\varphi}^* \tau^{\frac{\sigma}{\sigma-1}}$ where $\hat{\varphi}^* = \frac{\sigma^{\frac{\sigma}{\sigma-1}}}{\sigma-1} \left[\frac{wf}{P^\sigma Q}\right]^{\frac{1}{\sigma-1}} w$. Recall $I(\hat{\varphi})$ and $O(\hat{\varphi})$ where I is the cumulative input/labor share in the domestic market, and O is the cumulative sales share in the domestic market.

$$I(\hat{\varphi}) = \frac{\int \int_0^{\hat{\varphi}\tau^{\frac{\sigma}{\sigma-1}}} \varphi^{\sigma-1}\tau^{-\sigma} g(\varphi,\tau) d\varphi d\tau}{\int \int_0^{\inf} \varphi^{\sigma-1}\tau^{-\sigma} g(\varphi,\tau) d\varphi d\tau}$$

$$O(\hat{\varphi}) = \frac{\int \int_0^{\hat{\varphi}\tau^{\frac{\sigma}{\sigma-1}}} \varphi^{\sigma-1}\tau^{1-\sigma} g(\varphi,\tau) d\varphi d\tau}{\int \int_0^{\inf} \varphi^{\sigma-1}\tau^{1-\sigma} g(\varphi,\tau) d\varphi d\tau}.$$

Let $i(\hat{\varphi}) = I'(\hat{\varphi})$ and $o(\hat{\varphi}) = O'(\hat{\varphi})$. The hazard functions γ_s and γ_λ are

$$\gamma_s = -\frac{d\ln(1 - I(\hat{\varphi}))}{d\ln\hat{\varphi}} = \frac{i(\hat{\varphi})}{1 - I(\hat{\varphi})},$$

$$\gamma_\lambda = -\frac{d\ln(1 - O(\hat{\varphi}))}{d\ln\hat{\varphi}} = \frac{o(\hat{\varphi})}{1 - O(\hat{\varphi})},$$

When $\frac{i(\hat{\varphi})}{o(\hat{\varphi})}$ increases with $\hat{\varphi}$, i.e. I is likelihood ratio dominates O, then

$$\frac{1 - I(\hat{\varphi})}{i(\hat{\varphi})} = \int_{\hat{\varphi}} \frac{i(\hat{\varphi}')}{i(\hat{\varphi})} d\hat{\varphi}' \geq \int_{\hat{\varphi}} \frac{o(\hat{\varphi}')}{o(\hat{\varphi})} d\hat{\varphi}' = \frac{1 - O(\hat{\varphi})}{o(\hat{\varphi})},$$

that is, $\gamma_s \leq \gamma_\lambda$.

Let $x = \log\varphi, y = \log\tau$, then $x = \hat{\varphi} + \frac{\sigma}{\sigma-1}y$. Under joint-normal distribution of (x, y), define

$$V(\hat{\varphi}) \equiv \frac{i(\hat{\varphi})}{o(\hat{\varphi})} = \frac{\int \exp(\sigma x(\hat{\varphi}, y) - \sigma y) g(x(\hat{\varphi}, y), y) dy}{\int \exp(\sigma x(\hat{\varphi}, y) + (1-\sigma)y) g(x(\hat{\varphi}, y), y) dy}$$

where

$$g(x, y) = \exp\left[-\frac{1}{2(1-\rho^2)}\left(\frac{x^2}{\sigma_\varphi^2} + \frac{y^2}{\sigma_\tau^2} - \frac{2\rho xy}{\sigma_\varphi\sigma_\tau}\right)\right].$$

When $\sigma_\tau \geq \frac{\sigma-1}{\sigma}\rho\sigma_\varphi$, $V'(\hat{\varphi}) \geq 0$. Then the cumulative labor share distribution stochastically dominates the cumulative sales share distribution according to the likelihood ratio order, and the hazard functions satisfy $\gamma_s \leq \gamma_d$.

Furthermore,

$$\frac{d\ln\frac{1-I(\hat{\varphi})}{1-O(\hat{\varphi})}}{d\ln\hat{\varphi}} = \frac{d\ln(1 - I(\hat{\varphi}))}{d\ln\hat{\varphi}} - \frac{d\ln(1 - O(\hat{\varphi}))}{d\ln\hat{\varphi}} = -\gamma_s + \gamma_d \geq 0$$

then, it follows

$$\frac{1 - I(\hat{\varphi}_x^*)}{1 - I(\hat{\varphi}^*)} \geq \frac{1 - O(\hat{\varphi}_x^*)}{1 - O(\hat{\varphi}^*)}$$

and $S \leq \lambda$. Moving from a closed economy to an open economy, the reallocation term is always negative. Q.E.D. □

C Data: TFPR and TFPQ

We find large dispersions in measured TFPR in China, similar to the levels in HK for the year 1998 and 2007. Measured TFPR have come down over time, between 1998 and 2007, as evident in Table A-1. There is also greater dispersion in the average product of capital than there is in the average product of labor.

We next turn to investigating further what factors are systematically related to measured TFPR. Table A-2 reports the regression results of TFPR of a firm on a set of variables. The coefficient on firm TFPQ is large and significant; 1 percent increase in relative $TFPQ$ is associated with a 0.7 percent increase in relative TFPR. Moreover, more than half of the variation in TFPR is explained by TFPQ alone. The positive relationship is consistent with the predictions of our model as showing in the model regression. The same is true for the results on exporters: given TFPQ, firms must have lower taxes on average in order to export, and have a lower TFPR. TFPR differences are also systematic related to firm characteristics: state-owned enterprises and Foreign-owned firms are subject to lower TFPR on average, given TFPQ.

Table A-1: Dispersion of ARPK and ARPL

	1998	2001	2004	2007
std(ARPK)	1.348	1.306	1.241	1.185
std(ARPL)	1.184	1.039	0.940	0.923

Relationship between TFPR and TFPQ.

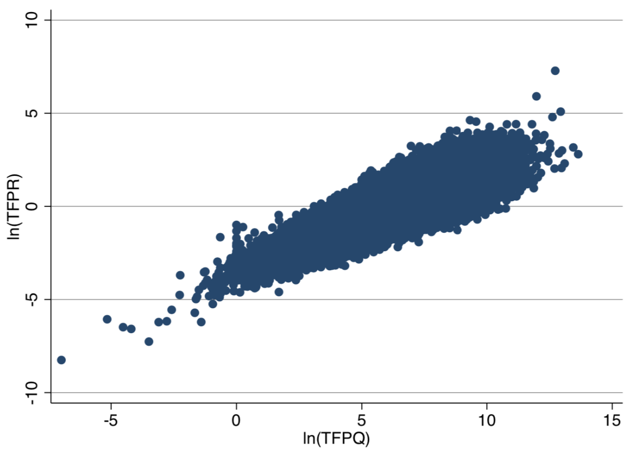
To show graphically, we next plot the relationship between $TFPQ$ and $TFPR$ in Figure A-1.

Table A-2: TFPR Regressions

VARIABLES	(1) ln($TFPR$)	(2) ln($TFPR$)	(3) ln($TFPR$)	(4) ln($TFPR$)	(5) ln($TFPR$)	(6) ln($TFPR$)
ln($TFPQ$)	0.574***	0.630***	0.635***	0.635***	0.635***	0.639***
	(243.5)	(235.9)	(243.2)	(241.6)	(248.4)	(261.6)
Age				-0.00165***	-0.00163***	-0.00148***
				(-9.736)	(-10.10)	(-10.05)
SOE					-0.100***	-0.0930***
					(-4.577)	(-4.481)
Foreign owned					-0.230***	-0.156***
					(-25.96)	(-24.60)
Exporters						-0.213***
						(-24.96)
Constant	-3.502***	-3.296***	-3.236***	-3.209***	-3.131***	-3.129***
	(-243.5)	(-106.2)	(-89.23)	(-87.12)	(-75.08)	(-77.04)
Observations	1,587,629	1,587,629	1,479,528	1,478,648	1,478,648	1,478,648
R-squared	0.739	0.812	0.822	0.823	0.831	0.837
Time FE	Yes	Yes	Yes	Yes	Yes	Yes
Industry FE	No	Yes	Yes	Yes	Yes	Yes
Location FE	NO	NO	YES	YES	YES	YES

Robust t-statistics in parentheses
*** $p<0.01$, ** $p<0.05$, * $p<0.1$

Figure A-1: Measured TFPR and TFPQ



D Equilibrium under Endogenous Markup

A closed-economy equilibrium consists of aggregate (P, Q, M) that satisfy:

$$M = \frac{\omega_e L}{Q \int_0^{\sigma^{\sigma/\varepsilon}} \int^{\hat{\tau}(\hat{q})} \left[\frac{\sigma}{\sigma - \hat{q}^{\frac{\varepsilon}{\sigma}}} \frac{\hat{q}}{\varphi}\right] g(\varphi(\tau, \hat{q}), \tau) \frac{d\varphi(\tau, \hat{q})}{dq} d\tau d\hat{q}}$$

$$Q \int_0^{\sigma^{\sigma/\varepsilon}} \int^{\hat{\tau}(\hat{q})} \left[\frac{\hat{q}^{\frac{\varepsilon}{\sigma}}}{\sigma - \hat{q}^{\frac{\varepsilon}{\sigma}}} \frac{\hat{q}}{\varphi}\right] g(\varphi(\tau, \hat{q}), \tau) \frac{d\varphi(\tau, \hat{q})}{dq} d\tau d\hat{q} = \omega_e f + f_e$$

$$\frac{M}{\omega_e} \int_0^{\sigma^{\sigma/\varepsilon}} \int^{\hat{\tau}(\hat{q})} \gamma(\hat{q})\, g(\varphi(\tau, \hat{q}), \tau) \frac{d\varphi(\tau, \hat{q})}{dq} d\tau d\hat{q} = 1,$$

where

$$\omega_e = \int_0^{\sigma^{\sigma/\varepsilon}} \int^{\hat{\tau}(\hat{q})} g(\varphi(\tau, \hat{q}), \tau) d\tau d\hat{q}$$

and

$$\gamma'\left(\frac{q}{Q}\right) = \frac{\sigma - 1}{\sigma} \exp\left(\frac{1 - (q/Q)^{\frac{\varepsilon}{\sigma}}}{\varepsilon}\right).$$

The open equilibrium consists of unknowns $(P, Q, M, P_f, Q_f, M_f, w_f)$ that satisfy:

$$\frac{\sigma}{\sigma - \hat{q}_x^{\frac{\varepsilon}{\sigma}}} \frac{w \tau_x \tau}{\varphi} = \gamma'(\hat{q}_x) P_f D_f$$

$$\pi_x = \left[\frac{\hat{q}_x^{\frac{\varepsilon}{\sigma}}}{\sigma - \hat{q}_x^{\frac{\varepsilon}{\sigma}}} \frac{\tau_x \hat{q}_x}{\varphi} Q_f - f_x\right] w,$$

where we get the zero profit cutoff. The free entry condition becomes:

$$\int \int_{\varphi^*(\tau)} \left[\frac{\hat{q}^{\frac{\varepsilon}{\sigma}}}{\sigma - \hat{q}^{\frac{\varepsilon}{\sigma}}} \frac{\hat{q}}{\varphi} Q - f\right] g(\varphi, \tau) d\tau d\varphi + \int \int_{\varphi_x^*(\tau)} \left[\frac{\hat{q}_x^{\frac{\varepsilon}{\sigma}}}{\sigma - \hat{q}_x^{\frac{\varepsilon}{\sigma}}} \frac{\tau_x \hat{q}_x}{\varphi} Q_f - f_x\right] g(\varphi_x, \tau) d\tau d\varphi = f_e \tag{A.19}$$

The labor market clearing condition is:

$$M = \frac{\omega_e L}{\int\int_{\varphi^*(\tau)} \left(\frac{\sigma}{\sigma - \hat{q}^{\frac{\varepsilon}{\sigma}}} \frac{\hat{q}}{\varphi} Q\right) g(\varphi,\tau) d\tau d\varphi + \int\int_{\varphi_x^*(\tau)} \left(\frac{\sigma}{\sigma - \hat{q}_x^{\frac{\varepsilon}{\sigma}}} \frac{\tau_x \hat{q}_x}{\varphi} Q_f\right) g(\varphi,\tau) d\tau d\varphi} \tag{A.20}$$

$$\left[\frac{M}{\omega_e} \int\int_{\varphi^*(\tau)} \gamma\left(\hat{q}\right) g(\varphi,\tau) d\tau d\hat{q} + \frac{M_f}{\omega_{ef}} \int\int_{\varphi_{xf}^*(\tau)} \gamma\left(\hat{q}_{xf}\right) g_f(\varphi,\tau) d\tau d\varphi\right] = 1 \tag{A.21}$$

For Foreign,

$$\begin{aligned} &\int\int_{\varphi_f^*(\tau)} \left[\frac{\hat{q}_f^{\frac{\varepsilon}{\sigma}}}{\sigma - \hat{q}_f^{\frac{\varepsilon}{\sigma}}} \frac{\hat{q}_f}{\varphi} Q_f - f\right] g_f(\varphi,\tau) d\tau d\varphi \\ &+ \int\int_{\varphi_{xf}^*(\tau)} \left[\frac{\hat{q}_{xf}^{\frac{\varepsilon}{\sigma}}}{\sigma - \hat{q}_{xf}^{\frac{\varepsilon}{\sigma}}} \frac{\tau_x \hat{q}_{xf}}{\varphi} Q - f_x\right] g_f(\varphi,\tau) d\tau d\varphi = f_{ef} \end{aligned} \tag{A.22}$$

$$M_f = \frac{\omega_e L_f}{\int\int_{\varphi_f(\tau)} \left(\frac{\sigma}{\sigma - \hat{q}_f^{\frac{\varepsilon}{\sigma}}} \frac{\hat{q}}{\varphi} Q_f\right) g_f(\varphi,\tau) d\tau d\varphi + \int\int_{\varphi_{xf}^*(\hat{q})} \left(\frac{\sigma}{\sigma - \hat{q}_{xf}^{\frac{\varepsilon}{\sigma}}} \frac{\tau_x \hat{q}_{xf}}{\varphi} Q\right) g_f(\varphi,\tau) d\tau d\hat{q}} \tag{A.23}$$

$$\left[\frac{M_f}{\omega_{ef}} \int\int_{\varphi_f^*(\tau)} \gamma\left(\hat{q_f}\right) g_f(\varphi,\tau) d\tau d\hat{q} + \frac{M}{\omega_e} \int\int_{\varphi_x^*(\tau)} \gamma\left(\hat{q}_x\right) g(\varphi,\tau) d\tau d\varphi\right] = 1 \tag{A.24}$$

Finally, the goods market clearing condition is:

$$\frac{M}{\omega_e} \int\int_{\varphi_x^*(\tau)} \left[\frac{\sigma}{\sigma - \hat{q}^{\frac{\varepsilon}{\sigma}}} \frac{w\hat{q}}{\varphi} Q_f\right] g(\varphi,\tau) d\tau d\varphi = \frac{M_f}{\omega_{ef}} \int\int_{\varphi_{xf}^*(\tau)} \left[\frac{\sigma}{\sigma - \hat{q}_{xf}^{\frac{\varepsilon}{\sigma}}} \frac{w_f \hat{q}_{xf}}{\varphi} Q\right] g_f(\varphi,\tau) d\tau d\varphi \tag{A.25}$$

E Regressions for measurement errors

Table A-3: Measured Marginal Products using First Differences vs TFPR

	(1)	(2)	(3)
VARIABLES	$\log(\frac{\triangle VA}{\triangle I})$	$\log(\frac{\triangle VA}{\triangle I})$	$\log(\frac{\triangle VA}{\triangle I})$
$\log(TFPR)$	0.718***	0.715***	0.718***
	(135.3)	(158.6)	(135.3)
Constant	1.410***	0.331***	1.410***
	(78.31)	(17.49)	(78.31)
Observations	624,659	624,699	624,659
R-squared	0.173	0.269	0.173
Time FE	Yes	Yes	Yes
Industry FE	Yes	Yes	Yes

Robust t-statistics in parentheses

*** p<0.01, ** p<0.05, * p<0.1

Specification (2) weights all the observations with the absolute value of composite input growth.

Specification (3) weights all the observations with the share of aggregate value added.

Table A-4: Estimate Measurement Error

VARIABLES	(1) $\triangle\widehat{VA}$	(2) $\triangle\widehat{VA}$	(3) $\triangle\widehat{VA}$
$\log(TFPR)$	0.0376***	0.0144***	0.0616***
	(22.62)	(9.170)	(16.07)
$[\log(TFPR)]^2$			-0.0128***
			(-6.110)
$[\log(TFPR)]^3$			0.00152***
			(4.008)
$\triangle\widehat{input}$	0.530***	0.523***	0.524***
	(34.58)	(33.03)	(31.13)
$\log(TFPR) \times \triangle\widehat{input}$	-0.0997***	-0.0954***	-0.0893***
	(-20.65)	(-19.16)	(-6.420)
$[\log(TFPR)]^2 \times \triangle\widehat{input}$			-0.00611
			(-0.919)
$[\log(TFPR)]^3 \times \triangle\widehat{input}$			0.00108
			(1.040)
Constant	-0.0207***	0.0551***	-0.0241***
	(-3.125)	(8.231)	(-3.592)
Observations	1,106,982	1,106,914	1,106,982
R-squared	0.044	0.042	0.044
Time FE	Yes	Yes	Yes
Industry FE	Yes	Yes	Yes

Robust t-statistics in parentheses

*** p<0.01, ** p<0.05, * p<0.1

Specification (2) weights all the observations with the share of aggregate value added.